



Università degli Studi di Milano
Scuola di Dottorato in Medicina Molecolare
Dipartimento di Scienze e Tecnologie Biomediche



Curriculum di *Genomica,*
Proteomica e Tecnologie Correlate

Ciclo XXIII

BIO -10

Anno Accademico 2009/2010

Dottorando: Arianna Gabrieli
Matricola: R07749

**STUDIO DI TECNOLOGIE DI AMPLIFICAZIONE
E GENOTIPIZZAZIONE DEL GENOMA SU
CAMPIONI DI DNA PROVENIENTI DA SANGUE
E DA CELLULE DELLA BOCCA PER
APPLICAZIONI IN AMBITO EPIDEMIOLOGICO.**

Direttore della Scuola: Ch.ma Prof.ssa Maria Luisa VILLA
Tutore: Prof. Cristina Battaglia
Correlatore: Dr.ssa Maria Carla Proverbio

INDICE

SOMMARIO.....	IV
ABSTRACT.....	V
1. INTRODUZIONE	- 1 -
1.1. Whole Genome Amplification	- 1 -
1.2. Cellule della bocca.....	- 4 -
1.3. Glutathione s-transferasi (GSTP1).....	- 6 -
2. SCOPO DEL PROGETTO.....	- 11 -
3. MATERIALI E METODI	- 12 -
3.1. Prelievo dei campioni.....	- 12 -
3.1.1. Estrazione del DNA da FTA Cards.....	- 13 -
3.1.2. Valutazione della stabilità nel tempo del DNA estratto dalle cellule della bocca.....	- 14 -
3.1.3. DNA estratto da sangue vs DNA estratto da cellule della bocca.....	- 14 -
3.2. Valutazione dell'utilizzo e affidabilità della WGA in biologia molecolare su campioni di DNA da sangue e da cellule della bocca.....	- 15 -
3.2.1. Confronto tra 2 kit commerciali	- 15 -
3.2.2. Purificazione dei prodotti amplificati con WGA.....	- 16 -
3.2.3. Valutazione della qualità, quantità dei prodotti WGA.....	- 16 -
3.2.4. Valutazione della fedeltà della WGA.....	- 17 -
3.3. Studio Epidemiologico: analisi del polimorfismo Ile105/Val105 del gene GSTP1	- 20 -
3.3.1. Analisi statistica	- 20 -
4. RISULTATI.....	- 21 -
4.1. Cellule della bocca e Sangue.....	- 21 -
4.1.1. Valutazione della qualità e quantità del DNA estratto da FTA Cards".	- 21 -

4.1.2.	<i>Valutazione della stabilità nel tempo del DNA estratto da cellule della bocca.....</i>	<i>- 22 -</i>
4.1.3.	<i>Utilizzo DNA estratto da sangue vs DNA estratto da cellule della bocca.....</i>	<i>- 22 -</i>
4.2.	<i>Valutazione dell'utilizzo e affidabilità della WGA in biologia molecolare su campioni di DNA da sangue e da cellule della bocca.....</i>	<i>- 24 -</i>
4.2.1.	<i>Confronto tra due kit commerciali</i>	<i>- 24 -</i>
4.2.2.	<i>Purificazione dei prodotti amplificati con WGA.....</i>	<i>- 25 -</i>
4.2.3.	<i>Valutazione della qualità e quantità dei prodotti WGA.....</i>	<i>- 25 -</i>
4.2.4.	<i>Valutazione della fedeltà della WGA</i>	<i>- 26 -</i>
4.3.	<i>Studio epidemiologico</i>	<i>- 36 -</i>
5.	DISCUSSIONE	- 37 -
6.	CONCLUSIONI	- 45 -
7.	RINGRAZIAMENTI	- 46 -
8.	BIBLIOGRAFIA	- 47 -
9.	PRODOTTI SCIENTIFICI DELL'ATTIVITA' DI DOTTORATO	- 51 -

SOMMARIO

Negli studi epidemiologici la quantità di materiale biologico disponibile rappresenta un fattore limitante. Molti studi utilizzano come campione biologico DNA ottenuto mediante prelievo venoso, tuttavia tale metodo di raccolta è invasivo soprattutto se i donatori sono bambini e persone anziane. L'utilizzo delle cellule della bocca rappresenta una risorsa alternativa, sebbene si ottenga DNA di scarsa qualità e quantità. Per incrementare la quantità di DNA estratto dalle cellule della bocca, si può utilizzare la "Whole Genome Amplification". Lo scopo del mio progetto di dottorato è stato quello di mettere a punto una metodica di estrazione di DNA da cellule della bocca e di studiare tecnologie di amplificazione e successiva genotipizzazione di DNA estratto da sangue e da cellule della bocca. La fedeltà della WGA è stata valutata con diverse tecniche di biologia molecolare: sequenziamento diretto, genotipizzazione mediante saggi di discriminazione allelica, microsatelliti, e studi "genome wide". Dalle nostre analisi è emerso che la WGA può essere impiegata per incrementare la quantità di materiale biologico di partenza, tuttavia presenta alcuni limiti, ad esempio nel sequenziamento diretto e nell'analisi con i microsatelliti, in alcuni casi, può determinare una perdita dell'informazione genetica. In seguito ai risultati ottenuti sull'utilizzo del DNA da cellule della bocca e della WGA abbiamo genotipizzato mediante Real Time il polimorfismo Ile105/Val105 del gene GSTP1 su 103 individui dell'area milanese. Lo studio delle frequenze alleliche di tale polimorfismo del gene GSTP1 è parte di un progetto, volto a valutare se in soggetti affetti da patologie respiratorie esiste un'interazione tra la predisposizione genetica individuale e l'esposizione a un comune agente esterno derivante dall'inquinamento urbano. Le frequenze genotipiche ottenute nella nostra popolazione non sono significativamente differenti da quelle della popolazione toscana (progetto HapMap), pertanto i nostri campioni potranno essere utilizzati come popolazione di riferimento in studi futuri. Inoltre abbiamo dimostrato che sia le cellule della bocca sia la WGA possono essere utilizzati in ambito epidemiologico per un'analisi di genotipizzazione mediante Real time PCR. La WGA può essere un metodo molto utile per aumentare la quantità di DNA e il DNA estratto da cellule della bocca può essere una valida risorsa per studi genetici.

ABSTRACT

In epidemiological studies the amount of biological material available is a limiting factor. Many studies use DNA as biological sample obtained by venipuncture, but this collection method is invasive especially if donors are children and the elderly. The use of mouth cells can be an alternative source, although you get DNA of poor quality and quantity. To increase the amount of DNA extracted from buccal cells, you can use the "Whole Genome Amplification". The aim of my PhD project was to develop a method to extract DNA from buccal cells and to study amplification technologies and subsequent genotyping of DNA extracted from blood and buccal cells. The accuracy of WGA was evaluated with different techniques of molecular biology and genotyping: direct sequencing, allelic discrimination assays, microsatellite genotyping and "genome wide analysis". Our analysis showed that the WGA can be used to increase the amount of starting biological material, however, it has some limitations, the fact that direct sequencing and analysis with microsatellites in some cases, may cause a loss of 'genetic information'. According to the data found using DNA from buccal cells and WGA, we have genotyped GSTP1 gene polymorphism Ile105/Val105 about 103 people in the Milan area through Real Time. The study of allele frequencies of this polymorphism in the GSTP1 gene is part of a project aiming to determine whether in patients with respiratory diseases there is an interaction between individual genetic predisposition and exposure to a common external agent coming from urban pollution. The genotypic frequencies obtained in our population were not significantly different from those of Tuscany population genotyped for the HapMap project, so our samples will be used as reference for future studies. Furthermore, we showed that both buccal cells and the WGA can be used in epidemiological analysis for genotyping through Real Time PCR. WGA may be a useful way to increase the amount of DNA; DNA extracted from buccal cells can be a valuable resource for genetic studies.

1. INTRODUZIONE

1.1. Whole Genome Amplification

Con il completamento del genoma umano e l'avvento di nuove tecnologie, oggi, è possibile effettuare studi "genome-wide" per identificare varianti genetiche associate ad alcune patologie. Un requisito fondamentale per gli studi "genome-wide" è che il DNA genomico e/o tumorale sia in quantità sufficiente. Sebbene molti studi di genotipizzazione si basino su una singola reazione di PCR, la quantità di DNA può essere insufficiente per studi di genotipizzazione che richiedono un numero maggiore di reazioni di amplificazione [1]. Da qui scaturisce la necessità di incrementare la quantità di DNA, sia per effettuare analisi su larga scala sia per conservare il DNA per studi genetici a lungo termine. Un approccio tradizionale consiste nell'immortalizzazione dei linfociti periferici mediante trasformazione con il virus di Epstein-Barr, però questa tecnica, piuttosto costosa, è applicabile quando si hanno a disposizione campioni di sangue appena prelevati e non DNA già estratto [1]. Una metodica economica e di semplice esecuzione, scoperta nel 1992, è la "Whole Genome Amplification" (WGA), basata su una reazione di amplificazione, la quale, a differenza di una reazione di PCR tradizionale che amplifica uno specifico locus genetico, permette un'amplificazione multi *locus* [2].

Esistono diverse tecniche di WGA, suddivise in due gruppi: metodi WGA basati su una reazione di PCR (DOP-PCR, PEP, LMP) e metodi WGA non basati su una reazione di PCR (MDA, TLAD). Nel 1992 [3] è stato scoperto il primo approccio WGA basato su una reazione di PCR: la DOP – PCR ("Degenerate Oligonucleotide Primed") in cui i *primers* oligonucleotidici parzialmente degenerati si legano a molti siti nel genoma durante diversi cicli di "annealing". Questo step di preamplificazione è seguito da un ciclo di amplificazione ad elevate temperature che permette solo ai frammenti che si sono legati ad una parte specifica di sequenza del *primer* di essere amplificati. Quest'approccio è particolarmente adoperato nel mappaggio del genoma.

La PEP [4] ("Primer Extension Preamplification") si basa sull'utilizzo di *primers* casuali di 15 basi che sono in grado di sintetizzare un frammento di DNA in circa 50 cicli di estensione del *primer*. Dato che non tutte le sequenze di DNA possono essere amplificate mediante PCR questa metodica è in grado di copiare il 78% del genoma da una singola cellula aploide almeno 30 volte, a causa della loro natura ripetitiva. Il protocollo basato sulla PEP è stato migliorato utilizzando un enzima a elevata processività e modificando alcune condizioni di PCR ("improved (I)-PEP"). Infine un approccio WGA basato su una reazione di PCR è la LMP "Ligation mediated PCR" [5], che implica il legame di una sequenza adattatore al DNA per facilitare la reazione di amplificazione. Tale metodica, a differenza della DOP-PCR e della PEP, ha il vantaggio di impiegare un solo *primer* per la reazione di

amplificazione. Tuttavia non esistono in letteratura lavori che utilizzano la LMP-PCR e ne valutano i vantaggi rispetto ad altri metodi.

Queste tecniche di WGA presentano dei difetti perché possono causare una maggiore amplificazione o una perdita di alcune regioni del genoma, sono soggette a una maggiore percentuale di errore di amplificazione e inoltre amplificano frammenti molto corti da 200 a 2000 paia di basi (bp) [6] (**Tabella 1**). Infatti, Dean e altri ricercatori [7] mediante saggio TaqMan hanno individuato un errore di amplificazione di circa 10^4 volte per la PEP e 10^6 per la DOP. Per queste ragioni molti ricercatori hanno preferito utilizzare approcci WGA non basati su una reazione di PCR, che assicurano una copertura completa del genoma e non introducono errori di amplificazione conservando l'informazione contenuta nel DNA [2]. Nel 2002 Iscoe [8] ha dimostrato l'utilizzo della TLAD ("T7-based linear amplification") mediante l'analisi di espressione dell'mRNA. Questa tecnica prevede la digestione del DNA con l'endonucleasi di restrizione AluI e la ligazione di un *primer* oligonucleotidico contenente un promotore T7 all'estremità 5' e un tratto di polyA all'estremità 3' del filamento di DNA. Tale metodica è poco usata poiché è molto costosa e richiede molto tempo.

Metodo	MDA^a	PEP^b	DOP^c
Resa del DNA (per 100 µl di reazione)	80 µg	40 µg	1-6 µg
Lunghezza dei prodotti di DNA (bp)	>10 Kb	100-1000	100-1000
Percentuale di errore della DNA polimerasi	< 10^{-6} nucleotidi	3×10^{-4} nucleotidi	3×10^{-4} nucleotidi

Tabella 1. Caratteristiche dei metodi di WGA: a_MDA (Multiple Displacement Amplification); b_PEP (Primer Extension Preamplification); c_DOP (Degenerate Oligonucleotide Primed).

Nel 2002 Dean ed altri ricercatori [7] hanno ideato un approccio WGA basato su un processo di amplificazione mediante spostamento del filamento ("Multiple Displacement Amplification": MDA) applicato al DNA genomico. Il processo di MDA si basa su un meccanismo di replicazione a circolo rotante (Rolling Circle Amplification, RCA.) [9] attraverso cui le molecole di DNA circolare si replicano come i virus e i plasmidi. Questa metodica prevede l'utilizzo di *primers* esamerici che hanno le estremità 3' tiofosfato modificate per proteggere i *primers* dalla degradazione dell'attività esonucleasica 3'-5' della DNA polimerasi, determinando un basso accumulo di mutazioni. Nella MDA è utilizzata la DNA polimerasi *Phi29*, derivata dal batteriofago *Bacillus Subtilis*, la quale presenta un'elevata processività, determinata dalla capacità di aggiungere 70.000 nucleotidi ogni volta che si lega ai *primers*, e dalla sua attività di spostamento del filamento [7]. In breve, il processo

consiste nell'utilizzo di *primers* esamerici che si legano al DNA in siti multipli che rappresentano i siti d'inizio della replicazione del DNA mediata dalla DNA Polimerasi. La replicazione procede lungo il filamento di DNA e i filamenti di DNA neo sintetizzati sono spostati e adoperati come stampo per nuovi eventi di replicazione, creando una struttura ramificata, come riportato in **Figura 1**. Il processo MDA genera prodotti di lunghezza > 10 kb con un contenuto di GC superiore all'80% (**Tabella 1**).

Questa metodica è impiegata per ottenere DNA amplificabile a partire da una minima quantità di campione biotico, sangue o cellule della bocca [10]. È impiegata, inoltre, per amplificare DNA plasmidico da impiegare in analisi di sequenziamento e serve come metodo di preparazione di campioni di DNA sostituendo i vecchi metodi di preparazione di DNA plasmidico da colonie o da culture [11]. Dato che l'MDA è utilizzata per ottenere DNA genomico per studi di genotipizzazione, se il DNA disponibile è di alta qualità, tale metodica copia entrambi gli alleli di un *locus* diploide [12].

In commercio sono presenti due kit basati sull'amplificazione WGA mediante MDA: *Repli-g Mini* kit (Qiagen) e *GenomiPhi* kit (GE Healthcare). Entrambi i kit prevedono una fase di denaturazione, mediante calore, del DNA e successiva reazione di amplificazione a 30°C in presenza di dNTPs, soluzione tampone e DNA polimerasi.

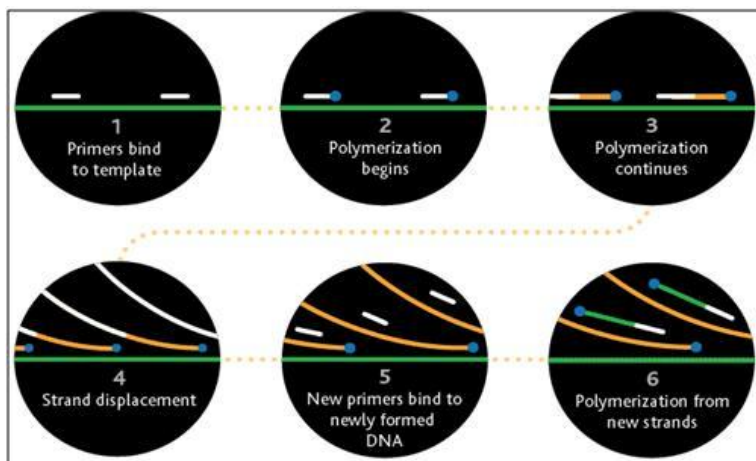


Figura 1. Processo di amplificazione basato sull'MDA.

1.2. Cellule della bocca

In ambito epidemiologico il sangue rappresenta il materiale biologico prediletto per ottenere DNA genomico in elevata quantità e di ottima qualità. Tuttavia bisogna considerare che il prelievo venoso presenta molti limiti: è necessario che sia effettuato da persone qualificate, può comportare infezioni, il materiale raccolto deve essere adeguatamente stoccato e trasportato in condizioni specifiche per evitare la degradazione del campione e nel caso di rottura delle provette contaminazioni con materiale potenzialmente infetto. Inoltre la procedura può essere molto invasiva quando il prelievo di sangue è eseguito in bambini o persone anziane. Molti epidemiologi hanno pensato di utilizzare materiali biologici alternativi come le cellule della bocca e la saliva. La loro raccolta è meno invasiva, più accessibile soprattutto a donatori anziani e in età pediatrica e inoltre permette di effettuare la raccolta in individui presenti in diverse aree geografiche del pianeta senza problemi di trasporto e di contaminazione con materiale pericoloso e infettivo [13]. In letteratura sono riportati diversi metodi di raccolta delle cellule della bocca: lavaggi della bocca, tamponi orali, raccolta su cartine da filtro [14]. Garcia Closas nel 2001 ha confrontato due metodi di raccolta: i lavaggi della bocca e i “citobrushes” [15], affermando che con i lavaggi della bocca si ottiene una maggiore quantità di DNA. Tuttavia il campionamento tramite i lavaggi della bocca implica che il donatore debba utilizzare una soluzione contenente alcool che, quindi, può recare una spiacevole sensazione di bruciore in bocca. Alternativamente il “citobrush” comporta lo sfregamento della mucosa orale e, quindi, allo stesso modo può recare una sensazione di fastidio. Inoltre tali metodi presentano un problema legato al trasporto, poiché il liquido contenuto in un apposito contenitore può versarsi o disperdersi durante il trasporto e inoltre richiedono spazi adeguati per lo stoccaggio e la conservazione.

Un altro metodo per la raccolta delle cellule della bocca si basa sull'utilizzo delle cartine da filtro opportunamente trattate [16]: FTA® cards (“Flinder Technology Associates”) che costituiscono una tecnologia innovativa brevettata dalla Whatman per la raccolta, il trasporto, la purificazione e l'archiviazione di DNA e RNA. Tali cartine contengono sostanze chimiche per la lisi delle cellule, la denaturazione delle proteine e proteggono gli acidi nucleici da nucleasi, ossidazione e danni da UV; inoltre inattivano rapidamente i microrganismi, compresi quelli patogeni veicolati dal sangue. Le applicazioni sull'uso delle FTA cards comprendono l'analisi medico legale, la medicina trasfusionale, ad esempio sono utilizzate per la raccolta di sangue, dai naturalisti per la raccolta di DNA degli uccelli per determinare i modelli di volo di alcune specie; per la raccolta di DNA di piante, batteri, plasmidi, microrganismi, particelle virali, cellule in cultura ecc. Per la raccolta di campioni incolori, come le cellule della bocca, è consigliabile utilizzare le FTA cards “Indicating”, le quali cambiano colore da rosa a bianco quando è applicato il

campione. Il campione è raccolto con un apposito applicatore sterile trasferito sulla cartina, dove rimane stabile per anni a temperatura ambiente [16] (**Figura 2**).



Figura 2. Cartine da filtro per la raccolta delle cellule della bocca ("Flinder Technology Associates": FTA card) e applicatore.

Le FTA card presentano vantaggi e svantaggi, riassunti nella tabella sottostante:

Caratteristiche	Vantaggi	Svantaggi
Raccolta	Facile da effettuare Non invasiva Limitata esposizione ai patogeni veicolati dal sangue Non necessita la presenza di personale medico	nessuno
Trasporto	Nessun rischio di versamento di liquidi	nessuno
Conservazione	Non richiede nessun processo particolare Conservazione a temperatura ambiente per circa 2 anni	Richiede condizioni adeguate di temperatura e pH.
Estrazione del DNA	Veloce e facile Processo non costoso	Scarsa qualità/quantità
Resa del DNA	~ 2 µg di DNA umano Possibilità di aumentare la resa di ~ 20µg mediante la Whole Genome Amplification.	Alcune volte non sufficiente per alcuni saggi

Tabella 2. Vantaggi e Svantaggi sull'utilizzo delle FTA Cards.

Dalle cellule della bocca, raccolte tramite cartine, si ottiene una quantità di DNA molto esigua, sufficiente per poche reazioni. Non ci sono molti studi che spiegano quali possono essere i fattori che influenzano la quantità e la purezza del DNA estratto. Nel 2009 Wijer [17] ed altri ricercatori hanno valutato se esistono fattori che possono limitare la quantità e la purezza del DNA estratto dalle cellule della bocca ed hanno osservato che la resa del DNA estratto dalle cellule della bocca degli uomini è maggiore di quello delle donne. Questa differenza tra uomini e donne può essere dovuto solo ad una differenza nella modalità di raccolta del campione. Inoltre non hanno evidenziato nessuna associazione tra la resa del DNA e l'età degli individui, in disaccordo con altri ricercatori che, precedentemente, avevano osservato una correlazione positiva tra l'età degli individui e la quantità di DNA [18]. La scarsa qualità del DNA estratto dalle FTA card, è dovuta probabilmente alla contaminazione con carboidrati, peptidi, fenoli, sali ed altri composti aromatici e questo inficia la riuscita di alcuni saggi [18,19]. Inoltre, come riportato dalla letteratura [16], un altro elemento presente nel DNA estratto da cellule della bocca o saliva che può influenzare la cattiva qualità del DNA estratto e la riuscita di alcuni saggi è la contaminazione del campione con DNA batterico, anche se ancora non è chiaro come il DNA batterico possa interferire con il DNA del soggetto in esame.

1.3. *Glutathione S-transferasi (GSTP1)*

Le glutathione S transferasi (GST) costituiscono una superfamiglia di proteine, note per la loro capacità di catalizzare l'attacco nucleofilo del glutathione ridotto (GSH) con atomi carenti di densità elettronica, presenti in vari xenobiotici, metaboliti di fase I e in alcuni prodotti del metabolismo cellulare. Sulla base della loro localizzazione subcellulare, le GST degli eucarioti possono essere suddivise in citosoliche, mitocondriali e microsomiali. Le GST citosoliche e mitocondriali sono proteine solubili e sono molto simili nella loro struttura quaternaria. Le GST microsomiali (MAPEG: "membrane –associated proteins in eicosanoid and glutathione metabolism") sono coinvolte nella biosintesi di prostanoidi e leucotrieni, molecole segnale endogene di natura lipidica [20]. Le GST citosoliche umane (hGST) sono molto polimorfiche e sono suddivise in sette classi, alfa, mu, pi, sigma, theta, zeta, omega, contenenti ciascuna diversi geni derivanti da eventi di duplicazione genica. Nella classificazione più recente il nome della sottofamiglia di ogni GST è abbreviato usando la prima lettera romana del nome della classe, seguita da un numero indicante il singolo prodotto genico. A tale numero può seguire la lettera indicante la variante allelica.

Le glutathione transferasi giocano un ruolo fondamentale nell'ambito di un meccanismo di difesa cellulare integrato in risposta allo stress chimico e ossidativo [21]. Precisamente, tali enzimi agiscono nella cosiddetta fase II dopo l'intervento del sistema del citocromo P-450, il principale responsabile delle reazioni di fase I, tra le quali la più importante sembra essere quella di ossigenazione. In questo

modo la fase I produce un metabolita più solubile in acqua e meno tossico che può essere escreto direttamente, ma che spesso è utilizzato come substrato per le reazioni della fase II. Queste ultime prevedono la coniugazione del metabolita attivato a un composto endogeno polare come il glutatione (GSH), l'acido glucuronico o la glicina. Nella maggioranza delle specie, la reazione che avviene più frequentemente nella fase II è rappresentata dalla coniugazione con il GSH, catalizzata dalle GST [20]. Non sempre, però, i composti substrato delle GST richiedono l'attivazione da parte del citocromo P-450; infatti, alcuni sono attivati dal ciclo ossigenasi, altri dall'interazione con i radicali liberi, altri ancora possono derivare direttamente dall'ambiente esterno, ad esempio in seguito ai processi di combustione. In ogni caso, le GST realizzano la loro funzione di protezione catalizzando l'attacco nucleofilo del GSH ridotto, sotto forma di anione tiolato (GS-TP1).

I geni *GST* sono molto polimorfici. Le varianti alleliche osservate originano da mutazioni (sostituzioni nucleotidiche), o, meno frequentemente, da delezioni.

La classe alfa delle GST umane comprende 5 geni e 7 pseudogeni, i quali formano un cluster localizzato sul cromosoma 6p12.1-6p12.2. I risultati epidemiologici mostrano che l'espressione aberrante di *GST* alfa è collegata all'aumento del rischio di sviluppare il cancro colo rettale, il cancro ovarico e il carcinoma renale a cellule chiare [22].

La classe delle *GST mu* comprende 5 geni localizzati in cluster sul cromosoma 1p13. Del gene *GSTM1* sono noti quattro alleli funzionali [23].

La classe delle *GST Theta* comprende 2 geni, entrambi localizzati sul cromosoma 22q11. Grazie a studi recenti è emerso che il gene *GSTT* influenza il rischio d'insorgenza di neoplasia e la risposta alla chemioterapia antitumorale [24].

La classe *GST Pi* è codificata da un singolo gene, localizzato sul cromosoma 11q13. La proteina GSTP è espressa, ad eccezione del fegato, in tutti i tessuti dell'organismo umano. La proteina GSTP1, attiva nel citosol, consiste di due subunità che formano un dimero. La proteina GSTP1 è sovra espressa in un'ampia gamma di tumori umani, inclusi carcinomi ovarici, polmonari, della mammella, del colon, del pancreas [25]. Del gene *GSTP1* sono note quattro differenti varianti alleliche: *GSTP1*A*, *GSTP1*B*, *GSTP1*C* e *GSTP1*D*. Questi alleli derivano da due transizioni localizzate nei codoni 105 (I105V) e 114 (A114V).

Le GST di classe Pi possiedono funzioni enzimatiche e non enzimatiche e sono coinvolte in molteplici processi cellulari come il metabolismo di fase II, la risposta allo stress ossidativo, la proliferazione cellulare, l'apoptosi, la progressione tumorale, e la resistenza ai farmaci. [20]

Il *GSTP1* è coinvolto nel processo di detossificazione da sostanze chimiche presenti nelle particelle di scarico diesel e costituisce un sistema di protezione intracellulare contro la formazione di specie reattive dell'ossigeno (ROS) pericolose, poiché è in grado di detossificare i composti elettrofili reattivi come gli idrocarburi aromatici policiclici (PAHs) [26]. Numerose ricerche dimostrano che i

lavoratori esposti durante la loro vita professionale ai fumi dei diesel hanno maggiore probabilità di contrarre il cancro ai polmoni rispetto alla popolazione in generale [26]. I fumi dei diesel inquinano l'aria che respiriamo ed è impossibile evitare di esporsi ad essi. Ovviamente le persone che vivono in ambiente urbano o industriale, dove circolano in gran numero mezzi pesanti, sono maggiormente esposte all'azione di questi inquinanti. I fumi dei motori diesel sono costituiti da particelle ultrafini (di spessore cinque volte inferiore al diametro di un capello) e penetrano profondamente nei polmoni. Recentemente è stato dimostrato che le sostanze assorbite sulle particelle entrano nel circolo sanguigno contribuendo a creare seri problemi per la salute. I fumi di scarico possono irritare gli occhi, il naso, la gola e i polmoni, possono provocare tosse, infiammazione ai polmoni aggravando i sintomi dell'insufficienza respiratoria cronica e aumentando la frequenza e l'intensità degli attacchi di asma [26].

L'esatto meccanismo con cui gli inquinanti presenti nell'aria possono provocare danni alla salute non è ancora chiaro. Tuttavia esistono delle varianti individuali che creano una diversa risposta agli effetti dell'inquinamento sulla salute, per cui l'individuazione di sottogruppi di popolazioni particolarmente suscettibili agli effetti dell'aria inquinata può essere un importante obiettivo: si possono identificare i soggetti a rischio sui quali applicare strategie di prevenzione. Tra i fattori esaminati come determinanti di questa variabilità, tra cui condizioni climatiche, dose e tempo di esposizione all'aria inquinata, età, sesso, dieta, presenza di malattie polmonari, un ruolo molto importante è svolto dai fattori genetici. In particolare tutti i processi coinvolti nella risposta all'inquinamento, quali stress ossidativo, infiammazioni, sono geneticamente regolati [27].

Recenti studi hanno messo in evidenza un'interazione tra l'esposizione ad agenti inquinanti e la variante Ile105/Val105 del gene *GSTP1* [28], in particolare la variante isoleucina (Ile)/valina (Val) nella posizione amminoacidica 105 è associata con l'incidenza di patologie allergiche [27,28]. Molti autori hanno studiato la distribuzione dei genotipi di tale polimorfismo in individui asmatici o allergici. Ad esempio Millard [29] ha individuato che individui con genotipo Val/Val presentano un basso rischio di sviluppare asma o allergia rispetto a individui portatori del genotipo Ile/Ile. Contrariamente Tamer [30] ha scoperto che individui con genotipo Val/Val presentano un maggiore rischio di sviluppare patologie allergiche. Questi dati contrastanti presenti in letteratura sono dovuti al fatto che sicuramente l'incidenza della patologia non è dovuta esclusivamente al polimorfismo, ma intervengono altri fattori quali l'esposizione ad aree a maggiore o minore inquinamento, il fumo, il sesso, l'età.

Per studiare l'interazione tra le varianti genetiche e l'inquinamento ambientale sono utilizzati due approcci: gli studi di linkage e gli studi di associazione. Fino ad oggi gli studi di linkage genetico sugli effetti dell'inquinamento sono stati eseguiti su modelli murini resi resistenti o suscettibili a determinati agenti inquinanti [26, 31]. La difficoltà di fare analisi di linkage sull'uomo nasce dall'incapacità di individuare

gli effetti sfavorevoli legati all'esposizione agli agenti inquinanti in diverse generazioni. Diversamente gli studi di associazione genetica hanno messo in relazione gli effetti dannosi degli agenti inquinanti tra individui con diversi genotipi in alcuni geni di rilevanza biologica [26]. I geni candidati devono codificare proteine coinvolte nei processi biologici legati alla risposta all'inquinamento ambientale. In molti studi di associazione sull'inquinamento sono stati studiati i geni coinvolti nelle difese antiossidanti (*Glutathione S-transferasi: GST*) e nell'infiammazione (*Tumor Necrosis Factor: TNF- α*). Finora sono stati fatti studi genetici di linkage e di associazione analizzando gli effetti dei polimorfismi dei geni dello stress ossidativo sull'inquinamento. In particolare nei modelli murini sono stati eseguiti studi di linkage e di associazione sulla risposta agli effetti dell'esposizione all'ozono e al particolato ambientale [32]. L'ozono è un agente inquinante gassoso. Negli studi su volontari sani [33] si è visto che l'esposizione all'ozono causa un cambiamento nella funzione polmonare, un'infiltrazione dei neutrofili nelle vie aeree e aumento dei livelli dei mediatori dell'infiammazione (IL8-ICAM). Negli studi di linkage nei modelli murini, sono state individuate due regioni sui cromosomi 17 e 11 associati all'infiltrazione dei neutrofili nelle vie aeree, fenotipo causato dall'esposizione all'ozono [32]. I geni presenti in queste regioni codificano proteine coinvolte nell'infiammazione (*TNF- α*) e nello stress ossidativo (superossido dismutasi).

I cromosomi 17 e 11 sono anche associati ad un danno acuto al polmone, all'edema polmonare dopo esposizione ad alte concentrazioni di ozono. I geni candidati in queste regioni includono geni antiossidanti (xantina deidrogenasi) e geni anti-infiammatori (*TNF- α*). Negli esseri umani sono stati fatti studi di associazione in seguito all'esposizione all'ozono, in particolare sono stati studiati i polimorfismi in due enzimi metabolici associati agli effetti dell'esposizione all'ozono: un enzima che produce idrochinoni (*NQO1*) il quale può reagire con l'ozono formando le specie reattive dell'ossigeno, e un enzima di detossificazione che riduce lo stress ossidativo (*GSTM1*) [34]. In uno studio fatto su ventiquattro individui non fumatori esposti a concentrazioni elevate di ozono, si è visto che gli individui con genotipi *NQO1* Pro/Pro187 e *GSTM1* nullo presentano un maggior danno a carico dell'epitelio respiratorio rispetto ad individui che esprimono il gene. In molti studi genetici sull'asma è stato riscontrato un aumento del rischio di asma in soggetti esposti ad elevate concentrazioni di ozono [33]. In uno studio fatto su bambini, affetti da asma, è stata individuata un'associazione tra i polimorfismi nei geni dello stress ossidativo (*NQO1* e *GSTM1*) e lo sviluppo di asma [35, 36]. Inoltre un trial clinico ha messo in evidenza che i bambini affetti da asma con genotipo *GSTM1* nullo presentano una perdita della funzionalità polmonare, se non sono trattati con sostanze antiossidanti come le vitamine C ed E [37]. E' stato riscontrato anche un aumento della difficoltà respiratoria nei bambini affetti da asma con genotipi legati allo stress ossidativo (*GSTM1* nullo e *GSTP1* Ile/Ile) [25]. Gli obiettivi principali degli studi gene-ambiente sull'inquinamento ambientale mirano a valutare l'esposizione agli agenti inquinanti e i loro effetti, e la loro interazione con i

fattori di rischio genetici. Fino ad oggi, gli studi di associazione sugli esseri umani rappresentano una strategia opportuna, ma sarebbe auspicabile fare studi su tutto il genoma.

2. SCOPO DEL PROGETTO

Lo scopo del mio progetto di tesi è articolato nei seguenti punti:

Messa a punto della metodica di estrazione del DNA da cellule della bocca raccolte su carta da filtro ("FTA Cards").

➤ Campioni utilizzati:

10 campioni appaiati di sangue e cellule della bocca dello stesso donatore.

Verifica dell'utilizzo del DNA estratto da cellule della bocca mediante tecniche di biologia molecolare (sequenziamento diretto, saggi di discriminazione allelica, studi "genome wide").

➤ Campioni utilizzati:

10 campioni appaiati di sangue e cellule della bocca dello stesso donatore.

Valutazione dell'utilizzo e affidabilità della WGA in biologia molecolare su campioni di DNA da sangue e da cellule della bocca e valutazione della concordanza dei genotipi tra i campioni di DNA da sangue e di cellule della bocca prima e dopo amplificazione mediante WGA.

➤ Campioni utilizzati:

10 campioni di DNA da cellule della bocca e da sangue

19 campioni di DNA da sangue così suddivisi:

16 campioni di DNA d'individui impiegati come controllo già presenti nel nostro laboratorio per uno studio su una patologia odontoiatrica.

3 campioni di DNA di pazienti affetti da iperinsulinismo congenito (presenti nel nostro laboratorio per uno studio su famiglie di pazienti affetti da una patologia congenita).

Studio epidemiologico.

Raccolta ed estrazione di campioni di cellule della bocca da individui sani residenti nell'area milanese per la valutazione delle frequenze genotipiche del polimorfismo Ile105/Val105 del gene *GSTP1*.

➤ Campioni utilizzati

103 campioni di cellule della bocca prelevati da individui sani residenti nell'area milanese.

Lo schema della procedura eseguita è rappresentato in figura 12.

3. MATERIALI E METODI

3.1. *Prelievo dei campioni*

Sangue

I campioni di sangue periferico sono stati prelevati in provette contenenti EDTA, previa richiesta di consenso informato e congelati a -20°C prima dell'estrazione del DNA.

Per garantire l'anonimato, i campioni di sangue sono stati etichettati con apposite sigle a seconda del progetto per il quale sono stati reclutati:

I campioni dei 10 soggetti che hanno donato anche le cellule della bocca sono stati identificati mediante la lettera N seguito da un numero progressivo.

I 16 campioni d'individui sani, reclutati per uno studio precedente in corso nel nostro laboratorio, sono stati identificati mediante l'iniziale del proprio nome e cognome seguito dalla data di nascita del donatore.

I 3 campioni di DNA d'individui affetti da iperinsulinismo, in cui erano state riscontrate delle mutazioni del gene *ABCC8*, sono stati identificati mediante la sigla HI (per indicare la patologia), seguita da un codice numerico (per indicare il nucleo familiare) e, in successione dalla lettera P (indicante il probando), o dalla lettera M se madre del probando o dalla lettera F (Padre del probando), dalla lettera B1/S1 (se primo fratello/sorella del probando).

Cellule della bocca

I campioni di cellule della bocca sono stati raccolti utilizzando cartine da filtro denominate "FTA cards". La procedura di raccolta, illustrata in **Figura 3**, consiste nel rimuovere un applicatore sterile dalla confezione e inserirlo in bocca sfregando uno dei lati del tampone in spugna all'interno della guancia per 30 secondi. La medesima operazione è ripetuta sull'altra guancia sfregando l'altro lato del tampone. Si fa scorrere l'applicatore lungo le gengive e sotto la lingua in modo da imbibirlo il più possibile di saliva. Dopo avere rimosso l'applicatore dalla bocca, si applica il campione sull'area circolare di colore rosa della cartina premendo la punta di spugna dell'applicatore e spremendola mediante un movimento di 90° per saturare completamente l'area. Si gira l'applicatore e si ripete con l'altro lato all'interno della stessa area circolare. Si lascia asciugare la cartina per circa 30 minuti a temperatura ambiente e quando la cartina è asciutta si può trasportare o conservare a temperatura ambiente fino al momento dell'estrazione del DNA.

Tutti i soggetti reclutati per il nostro studio hanno effettuato la raccolta in modo autonomo, possibilmente a digiuno, previa compilazione del modulo di consenso informato e di una scheda di anamnesi.

Al fine di garantire l'anonimato, i campioni raccolti sono stati identificati mediante l'assegnazione della lettera SN seguita dallo stesso numero progressivo attribuito al corrispondente campione di sangue.



Figura 3. Procedura di raccolta delle cellule della bocca mediante le FTA Cards

3.1.1. Estrazione del DNA da FTA Cards

Per l'estrazione del DNA dalle cellule della bocca sono stati tagliati piccoli dischetti di 3 mm di diametro dalla cartina con un apposito strumento. Al fine di ottenere una quantità adeguata di DNA sono state effettuate prove con un diverso numero di dischetti (5-10-20 dischetti). Il DNA è stato estratto dai dischetti con il kit Pure Link Genomic DNA (Invitrogen), che si basa su un metodo di estrazione su colonnine, utilizzando il protocollo "Blood Spots" riportato dalla ditta fornitrice del kit. Brevemente, i dischetti sono stati trasferiti in una provetta ed è stato aggiunto un tampone di digestione e la proteinasi K per la digestione delle proteine; la miscela è stata incubata a 55 °C e, in seguito, al lisato sono stati aggiunti l'RNasi A, per degradare l'RNA, un tampone di legame ed etanolo per favorire il legame del DNA alla membrana della colonnina nella quale è stato trasferito. Dopo lavaggi con tamponi al fine di rimuovere le impurità presenti, il DNA è stato eluito in H₂O e l'eluato conservato a -20°C.

Come riportato nel diagramma a flusso di figura 12, i campioni di DNA estratti sono stati sottoposti ai controlli qualitativi e quantitativi mediante lettura spettrofotometrica al Nanodrop, gel di agarosio 0.9% e PCR "multiplex" (DNA OK kit, Labogen). In particolare, la resa di DNA e la qualità sono state valutate mediante analisi spettrofotometrica utilizzando il Nanodrop ND 100 (Nanodrop Technologies Wilmington, DE), dove i rapporti A260/A280 e A260/A230 indicano il rapporto tra l'assorbanza degli acidi nucleici (A260) e l'assorbanza dell'RNA (A280) e delle proteine (A230) rispettivamente. Il rapporto ottimale che indica la buona qualità del DNA estratto deve essere ≥ 1.8 . Inoltre la qualità del DNA è stata valutata mediante il kit DNA OK (Labogen), il quale si basa su una reazione di PCR multipla in cui si ottengono ampliconi da 100 a 600 paia di basi se il DNA è perfettamente integro, se sono presenti meno di 6 bande il DNA è degradato. Il protocollo di amplificazione utilizzato è quello descritto dalla ditta produttrice. Infine, la qualità e la quantità del DNA sono state stimate mediante corsa elettroforetica su gel di agarosio 0.9%.

3.1.2. Valutazione della stabilità nel tempo del DNA estratto dalle cellule della bocca.

Al fine di valutare la stabilità del DNA nel tempo, i campioni di DNA sono stati amplificati mediante il Kit DNA OK (Labogen) al momento dell'estrazione, dopo 6 e 12 mesi dall'estrazione.

3.1.3. DNA estratto da sangue vs DNA estratto da cellule della bocca.

Sequenziamento diretto

L'esone 25 del gene "ATP-binding Cassette, Subfamily C, member 8" (ABCC8, GeneID:6833), è stato amplificato mediante PCR e poi sequenziato nei 10 campioni di DNA estratto da sangue e da cellule della bocca. I *primer* utilizzati sia per la reazione di PCR che per la reazione di sequenza sono Ex25F_5'AGCCACAGAGTGGGTCT3'

Ex25R_5'CAGGCAGAGAGTGATTTGGAG3'.

Le reazioni di PCR sono state eseguite in 30 µl di volume di reazione utilizzando 100 ng di DNA, PCR Buffer 10X (Roche), MgCl₂ 25 mM, dNTPs 2,5 mM, DNA polimerasi FAST START 5 U/ µl (Roche). Il protocollo termico consiste:

Denaturazione iniziale a 95°C per 4 min

Denaturazione a 95°C per 30 sec

"Annealing" a 51°C scelta per 30 sec

Estensione a 72°C per 45 sec

Estensione finale a 72°C per 8 min

} per 40 cicli.

I prodotti di PCR sono stati controllati mediante corsa su gel di agarosio 2,5%. I prodotti di PCR sono stati purificati con "ExoSAP" che utilizza due enzimi idrolitici: l'esonucleasi I e la fosfatasi alcalina, per rimuovere i dNTPs e i *primers*. In particolare l'esonucleasi I degrada i primers a singolo filamento rimasti e i filamenti di DNA a singolo filamento estranei. La fosfatasi alcalina idrolizza i dNTPs rimasti dalla reazione di PCR, che potrebbero interferire con la reazione di sequenza. I prodotti di PCR purificati sono stati sequenziati utilizzando il kit di sequenza "DYEnamic ET terminator Sequencing Premix" (Amersham). I prodotti di sequenza sono stati purificati per filtrazione con piastre SEQ95 (Millipore) e sottoposti a elettroforesi capillare su sequenziatore automatico a 96 capillari Megabace 1000 (Amersham). Le sequenze ottenute sono state analizzate mediante il software Sequence Analyzer Cimarron 3.12 (Amersham) e confrontate manualmente con le sequenze depositate in banca dati.

Saggi di discriminazione allelica mediante TaQMan Assay.

I campioni di DNA estratti da sangue e da cellule della bocca sono stati genotipizzati mediante il saggio di discriminazione allelica per il polimorfismo rs1695 (A/G) del gene *GSTP1* (GeneID:2950) e il polimorfismo rs 4678013 (G/T)

del gene calcium sensing receptor (CASR: GeneID.846). E' stato utilizzato il kit "TaqMan Assay" (Applied Biosystems) contenente i due *primers* "forward" e "reverse" e le due sonde per individuare i due alleli, ciascuna marcata con un fluoroforo in particolare il fluoroforo FAM all'estremità 5' per individuare l'allele 2 e il fluoroforo VIC all'estremità 5' per individuare l'allele 1. La reazione di PCR è stata eseguita in 15 µl di volume finale con 20 ng di DNA, master mix (Takara) 2X 7.5 µl, TaqMan Genotyping Assay 20X (Applied Biosystems) 0.75 µl. I campioni sono stati analizzati in duplicato: come controllo sono stati inseriti 3 campioni a genotipo noto. In Tabella 3 sono riportati i *primers*.

Gene	Codice Applied Biosystems	Sequenza primer [VIC/FAM] (5'-3')
CASR rs46780 13	C_8332561_20	CTGAGCTTACTTGGAAACCCTTTT[G/T] GGATCTGTCCAAAAGGAGTGAATGT
GSTP1 rs1695	C_3237198_20	CGTGGAGGACCTCCGCTGCAAATAC[A/G] TCTCCCTCATCTACACCAACTATGT

Tabella 3. Primer dei geni analizzati mediante saggi TaqMan.

Il protocollo termico consiste in:

Denaturazione iniziale a 95°C per 30 sec

Denaturazione a 95°C per 5 sec } per 45 cicli

"Annealing" a 60°C per 20 sec

Le chiamate genotipiche di ogni singolo campione sono state assegnate mediante il software iCycler iQ (Biorad).

3.2. Valutazione dell'utilizzo e affidabilità della WGA in biologia molecolare su campioni di DNA da sangue e da cellule della bocca

3.2.1. Confronto tra 2 kit commerciali

Protocollo di WGA utilizzando il kit Repli-G Mini kit (Qiagen.)

4 ng di DNA estratto da sangue e da cellule della bocca sono stati utilizzati per la reazione di amplificazione. Brevemente, 20 ng di template sono stati mescolati con 2,5 µl di buffer di denaturazione (D1). La miscela è stata incubata a temperatura ambiente per 3 minuti. Successivamente, alla miscela è stato aggiunto 2,5 µl di buffer di neutralizzazione (N1). 40 µl di mix costituita da buffer di reazione e DNA polimerasi sono stati aggiunti a 10 µl di DNA denaturato. I prodotti sono stati incubati a 30 °C per 16 ore. In seguito, per inattivare l'enzima, i campioni sono stati sottoposti a un ciclo di 65°C per 3 minuti. I campioni amplificati sono stati conservati in ghiaccio.

Protocollo di WGA utilizzando il kit Genomiphi (GE Healthcare.)

Circa 20 ng di DNA sono stati mescolati con 22.5 µl di “sample buffer”. La miscela è stata denaturata a 95°C per 3 minuti e immediatamente è stata posta in ghiaccio. 22.5 di µl “reaction buffer” e 2.5 µl di enzima sono stati aggiunti alla miscela, precedentemente posta in ghiaccio. Il prodotto ottenuto è stato incubato a 30°C per 4 ore. Per inattivare la DNA polimerasi, i campioni sono stati sottoposti a un ciclo di 65°C per 10 minuti.

Dopo amplificazione con WGA al codice identificativo di tutti i campioni è stata aggiunta la sigla GP.

3.2.2. Purificazione dei prodotti amplificati con WGA.

Per rimuovere i residui di dNTPs e di buffer i prodotti di WGA sono stati purificati con due differenti metodi.

Purificazione mediante la precipitazione con sodio acetato/EDTA.

A ciascun campione sono stati aggiunti 50 µl di H₂O, 1/10 del volume del tampone sodio acetato/EDTA e 100 µl di etanolo 100%. La miscela è stata incubata a -20°C per tutta la notte. Il giorno dopo la miscela è stata centrifugata a temperatura ambiente per 15 min a 12000 rpm; il surnatante è stato rimosso e il pellet è stato lavato con 100 µl di etanolo 70% e risospeso in 30 µl di H₂O.

Purificazione utilizzando il kit Qiaamp DNA Mini kit (Qiagen.)

Il campione è stato trasferito in una colonnina fornita dal kit, 50 µl di “lysis binding buffer” sono stati aggiunti al campione. La miscela è stata centrifugata a 12000 rpm per 1 min, il campione è stato lavato con il “washing buffer 1”, centrifugato a 12000 rpm per 1 min, e poi lavato con il “washing buffer 2” e centrifugato a massima velocità per 3 min, infine il campione è stato eluito in 200 µl di H₂O.

3.2.3. Valutazione della qualità, quantità dei prodotti WGA

PCR “multiplex” (DNA OK, Labogen)

I prodotti GP-DNA sono stati controllati qualitativamente e quantitativamente mediante il kit DNA OK, utilizzando il protocollo fornito dalla ditta.

Corsa elettroforetica su Gel di agarosio 0.9%.

I campioni GP-DNA sono stati caricati su gel di agarosio 0.9% insieme con il marcatore ad alto peso molecolare (Invitrogen) per valutare la presenza di frammenti >10000 bp.

Lettura spettrofotometrica mediante il Nanodrop-ND 100.

1 µl di campione è stato letto allo spettrofotometro, riportando la quantità di ciascun campione e il rapporto A260/280 (assorbanza degli acidi nucleici/assorbanza delle proteine), indice della qualità del campione.

Tutti i controlli di qualità dei prodotti GP-DNA sono stati ripetuti dopo 6 e 12 mesi.

3.2.4. Valutazione della fedeltà della WGA

Sequenziamento diretto

L'esone 25 del gene "ATP-binding Cassette, Subfamily C, member 8" (ABCC8, GeneID:6833), l'esone 2 del gene "Tumor Necrosi Factor" (TNF- α , GeneID:7124) e l'esone 5 del gene *Glutathione S-transferasi* (GSTP1, GeneID:2950) sono stati amplificati e sequenziati nei campioni GP-DNA da sangue e da cellule della bocca parallelamente ai rispettivi campioni di DNA genomico. Inoltre sono stati amplificati gli esoni 10,39, 37 del gene ABCC8 nei 3 campioni affetti da HI in cui sono state riscontrate mutazioni. Le reazioni di PCR sono state eseguite con le stesse modalità riportate precedentemente. I *primers* utilizzati per la PCR e la reazione di sequenza sono riportati nella tabella sottostante (**Tabella 4**). I prodotti di PCR, dopo essere stati controllati mediante corsa elettroforetica su gel di agarosio 2,5%, sono stati purificati mediante filtrazione con piastre MICRO PCR96 (Millipore) e sequenziati con gli stessi *primers* impiegati per la PCR utilizzando il kit ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction 3.1 (Applied Biosystems). I prodotti di sequenza sono stati purificati per filtrazione con piastre SEQ95 (Millipore) e sottoposte a elettroforesi capillare su sequenziatore automatico *ABI Prism 377 DNA Sequencer* (Applied Biosystems). Le sequenze ottenute sono state analizzate mediante il software Sequence Scanner v1.0 e confrontate manualmente con le sequenze depositate in banca dati.

Gene	Sequenza primers	TA*
ABCC8	Ex25F_5'AGCCACAGAGTGGGTCT3' Ex25R_5'CAGGCAGAGAGTGATTTGGAG3'	51°C
TNF- α	Ex2F_5'CTCAGGACTCAACACAGCTT3' Ex2R_5'CCACTACCGCTTCCTCCAGA3'	57°C
GSTP1	Ex5F_5'CACATCCTCTTCCCTCCTC3' Ex5R_5'CAAGCCACCTGAGGGGTAA3'	64°C

Tabella 4. Primers dei geni analizzati mediante sequenziamento diretto.

(*TA: temperatura di "annealing")

Analisi del microsatellite HUM-FIBRA (FGA)

Il marcatore tetra nucleotidico TTTC HUM-FIBRA (FGA) è stato analizzato su 16 campioni di DNA genomico estratto da sangue e sui rispettivi campioni GP-DNA. La ripetizione tetra nucleotidica è localizzata nel terzo introne del gene del fibrinogeno umano sul cromosoma 4. La PCR è stata eseguita in 15 μ l di reazione finale, utilizzando 20 ng di DNA, Buffer (1X) 1.5 μ l, dNTPs (2,5 mM) 1.5 μ l, MgCl₂ (2.5 mM) 1.5 μ l, Taq Gold Polimerasi (5 unità/l) 0.12 μ l, e il *primer forward*

(marcato con il fluoroforo HEX) 10 μ M e il *primer reverse* 10 μ M. Il protocollo termico prevede:

Denaturazione iniziale a 95°C per 12 minuti

“Melting” a 94°C per 15 sec

Estensione dei primers alla TA ottimale per 15 sec

Estensione finale a 72°C per 30 sec

per 10 cicli

“Melting” a 89°C per 15 sec

Estensione dei primers alla temperatura ottimale per 15 sec

Estensione finale a 72 °C per 30 sec.

per 20 cicli

I campioni amplificati sono stati caricati su ABI Prism 377 DNA Sequencer (Applied Biosystems) e i prodotti sono stati analizzati utilizzando il software Peak Scanner.

Saggi di discriminazione allelica mediante TaqMan Assay

26 campioni di DNA estratti da sangue e 10 campioni da cellule della bocca prima e dopo amplificazione con WGA sono stati genotipizzati mediante il saggio di discriminazione allelica per il polimorfismo rs1695 del gene GSTP1 e il polimorfismo rs 4678013 del gene calcium sensing receptor (CASR). Sono stati utilizzati i kit “TaqMan Assay” (Applied Biosystems), utilizzando il medesimo protocollo precedentemente riportato per confrontare i genotipi dei campioni di DNA da sangue vs DNA da cellule della bocca.

Le chiamate genotipiche di ogni singolo campione sono state assegnate mediante il software iCycler iQ (Biorad).

Genotipizzazione mediante Gene Chip Human Mapping 250K Nspl Array

○ *Protocollo Affymetrix Gene Chip Human Mapping 250K Nspl Array*

3 coppie di campioni di DNA estratti da sangue prima e dopo amplificazione mediante WGA (HI15M/HI15MGP, HI25M/HI25MGP, HI33F/HI33FGP) e un campione di sangue e cellule della bocca dello stesso individuo prima e dopo amplificazione con WGA (N5/N5GP, SN5GP) sono stati impiegati per l'analisi di SNPs su ampia scala utilizzando il kit GeneChip® Human Mapping 250K Nspl, seguendo il protocollo fornito dal kit (**Figura 4**). Il campione SN5 non ha una concentrazione di DNA sufficiente per analisi “genome wide” con la piattaforma Gene Chip Human Mapping 250K Nspl Array (50 ng/ μ l in un volume massimo di 5 microlitri), pertanto il campione di cellule della bocca dopo WGA (SN5GP) è stato analizzato con il suo corrispettivo campione di sangue.

Brevemente, 250 ng di DNA (prima e dopo amplificazione con WGA) è stato digerito con l'enzima di restrizione Nspl, legato ad adattatori sintetici e amplificato mediante PCR utilizzando primers universali che si legano alla sequenza adattatore. La lunghezza dei prodotti di PCR (da 250-1110 bp) è stata controllata mediante elettroforesi capillare (Agilent Bioanalyzer). Le reazioni di PCR di ogni campione sono state unite in un unico campione, e dopo purificazione i prodotti di PCR sono stati quantizzati allo spettrofotometro (NanoDrop ND-100) e 90 μ g di prodotto di PCR sono stati frammentati mediante la DNasi I a 37°C per 35 minuti. I

prodotti frammentati sono stati controllati mediante elettroforesi capillare. Successivamente, sono stati etichettati all'estremità con la deossinucleotide transferasi a 37°C per 2 ore e ibridati sull'array" per 8 ore a 49°C. Dopo ibridazione, gli "array" sono stati lavati e colorati mediante l'utilizzo della fluidica "Affymetrix Fluidics Station 450". Successivamente, l'immagine fluorescente dei "microarrays" è stata acquisita mediante lo scanner "Gene Chip® Scanner 3000" e analizzata con il software GYTYPE v4.0 (AFFYMETRIX) per quantizzare l'intensità di segnale corrispondente a ciascuno SNP. GeneChip® Human Mapping 250K array è in grado di genotipizzare 262,264 SNPs con una distanza media intermarcatore di 9.4 kb su un solo chip.

Il controllo di qualità del "chip", la normalizzazione dell'intensità della sonda e la creazione delle chiamate genotipiche, sono state eseguite utilizzando il software GeneChip® Genotyping (GYTYPE v4.1). L'algoritmo BRLMM (Robust Linear Model with Mahalanobis distance classifier) è stato impiegato per ottenere le chiamate genotipiche degli SNP array 250K (una SNP "call rate">95% è un parametro di buona qualità dello SNP array).

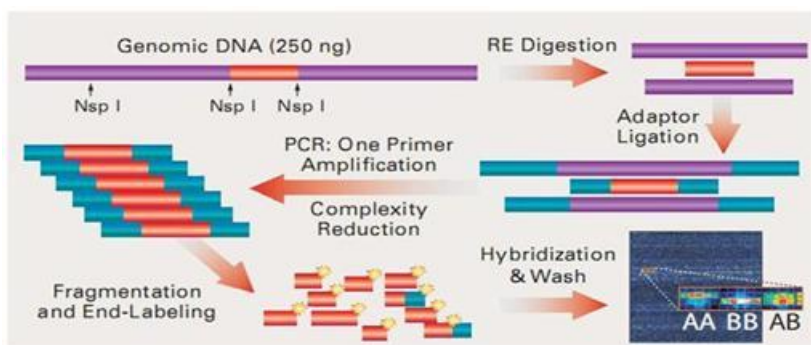


Figura 4. Rappresentazione schematica del protocollo Affymetrix GeneChip Human Mapping 250K NspI Arrays.

○ *Analisi delle alterazioni di "copy number" (CN) e perdita di eterozigotà (LOH).*

L'analisi di "copy number" (CN) e della perdita di eterozigotà (LOH) è stata effettuata utilizzando il software CNAG v3.0. Tale software utilizza un algoritmo di riferimento, denominato AsCNAR. Tale algoritmo è in grado di effettuare simultaneamente l'analisi di LOH e di CN in campioni appaiati e non appaiati utilizzando una raccolta di 48 campioni CEU (Centre d'Etude du Polymorphisme Humain collection), reclutati dal progetto HAP MAP, come controllo (pubblicati su: www.hapmap.org/downloads/raw_data/affy500k/). CNAG v3.0 utilizza i files risultanti dall'analisi di genotipizzazione (CHP) e i files d'intensità cellulare (CEL) esportati da GTYPE, convertendo l'intensità di segnale in un valore di "copy number" per ogni SNP. Per queste analisi, abbiamo utilizzato parametri di riferimento (probabilità di LOH > 30). Dopo la normalizzazione la compensazione e

la media è stata calcolata su una finestra di 10 SNPs consecutivi, all'interno di ogni finestra di analisi gli SNPs con un valore di "copy number" più alto e più basso vengono scartati per ridurre il rumore di fondo.

3.3. Studio Epidemiologico: analisi del polimorfismo Ile105/Val105 del gene *GSTP1*

Il polimorfismo Ile105/Val105 del gene *GSTP1*, è stato analizzato in 103 individui residenti nell'area milanese. Il campione di cellule della bocca è stato prelevato mediante le "FTA Cards". Ad ogni individuo è stato somministrato un modulo di consenso informato ed è stata chiesta di compilare una scheda di anamnesi. Sulla base dei dati raccolti, in **Tabella 5** sono riportati gli individui suddivisi in maschi e femmine, fumatori –non fumatori.

Campioni (n=103)	Sesso		Fumo			
	M	F	Non fumatori	Fumatori <10 sigarette	Fumatori >10 sigarette	Ex fumatori
	28	75	79	14	3	7

Tabella 5. Individui suddivisi per sesso, fumo

Il DNA è stato estratto come precedentemente riportato e successivamente è stato sottoposto ai seguenti passaggi:

- Controlli di qualità e quantità come riportato nel paragrafo 3.1.1.
- Amplificazione mediante il kit "GenomiPhi", utilizzando il protocollo descritto nel paragrafo 3.2.1.
- Purificazione dei prodotti amplificati con WGA come descritto nel paragrafo 3.2.2.
- Genotipizzazione dello SNP Ile105/Val 105 del gene *GSTP1* mediante saggio "TaqMan Assay" (Applied Biosystems) dei 103 campioni GP-DNA, seguendo il protocollo riportato nel paragrafo 3.2.4. I genotipi sono stati assegnati utilizzando il software dello strumento iCycler iQ (Biorad).

3.3.1. Analisi statistica

Le frequenze genotipiche del polimorfismo Ile/Val del gene *GSTP1* dei 103 campioni sono state confrontate con le frequenze genotipiche di una popolazione di riferimento costituita da 114 individui toscani reclutati per il progetto HapMap (<http://hapmap.ncbi.nlm.nih.gov/>) mediante il test statistico chi quadro (χ^2).

4. RISULTATI

4.1. Cellule della bocca e Sangue

4.1.1. Valutazione della qualità e quantità del DNA estratto da FTA Cards”.

La lettura spettrofotometrica mediante Nanodrop ND-100 del DNA estratto da 5, 10 e 20 dischetti di “FTA Cards” ha evidenziato che la resa migliore è stata ottenuta col maggior numero di dischetti impiegati per l'estrazione, come riportato in **Tabella 6**. Aumentando il numero dei dischetti la quantità non varia in modo significativo (dati non riportati).

Campione	5 dischetti	10 dischetti	20 dischetti
1	8.9 ng/μl	10 ng/μl	24 ng/μl
2	9 ng/μl	9 ng/μl	18 ng/μl

Tabella 6. Quantità di DNA ottenuta da 2 campioni di cellule della bocca raccolti mediante FTA Card variando il numero dei dischetti utilizzati per l'estrazione.

La concentrazione di DNA ottenuta dai 10 campioni di cellule della bocca utilizzati per la messa a punto della metodica di estrazione è compresa tra 8 ng/μl e 28 ng/μl con un valore di assorbanza A260/280 compreso tra 0.8 e 1.3 (**Tabella 7**); la concentrazione del DNA estratto dai 10 campioni di sangue varia tra 21 ng/μl e 60 ng/μl con un valore di A260/280 compreso tra 1,7 e 2 (dati non riportati).

Campione	DNA (ng/μl)(*)	A260/280	DNA totale (μg)(*)
SN1DNA	14	1.3	1.4
SN2DNA	28	0.4	2.8
SN3DNA	19	0.7	1.9
SN4DNA	8	0.6	0.8
SN5DNA	16	0.8	1.6
SN6DNA	8	0.7	0.8
SN7DNA	9	0.7	0.9
SN8DNA	9	0.9	0.9
SN9DNA	14	0.4	1.4
SN10DNA	10	0.8	1

Tabella 7. Concentrazione (ng/μl) e rapporto A260/280 ottenuti mediante lettura spettrofotometrica di 10 campioni di DNA estratti da cellule della bocca, () i campioni sono stati eluiti in 100 μl di H₂O.*

Il DNA estratto dalle cellule della bocca è stato anche controllato qualitativamente, mediante corsa su gel di agarosio allo 0.9 % dove è visibile solo uno "smear" (dato non riportato) e mediante il kit DNA OK (Labogen). Gli ampliconi ottenuti sono stati controllati mediante corsa su gel di agarosio all'1%. La presenza delle 6 bande di 100-200-300-400-500-600 paia di basi, indica la buona qualità del campione estratto (**Figura 5**).

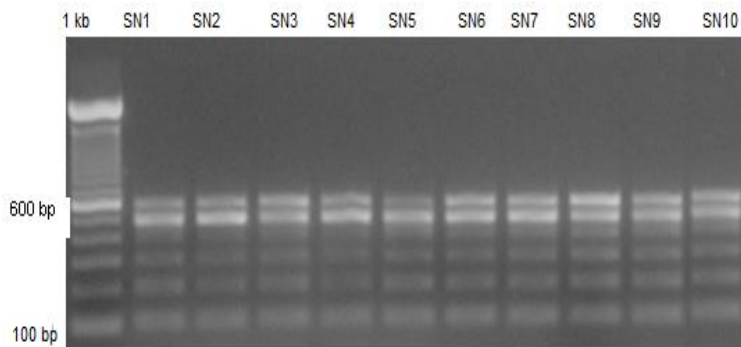


Figura 5. Gel di agarosio 1% degli ampliconi di 10 campioni di DNA estratti da cellule della bocca, ottenuti mediante PCR multiplex (DNA OK, Labogen), al momento dell'estrazione.

4.1.2. Valutazione della stabilità nel tempo del DNA estratto da cellule della bocca.

Il DNA estratto dai 10 campioni di cellule della bocca è stato controllato con DNA Ok dopo 6 mesi (dati non riportati) e dopo 12 mesi per valutare la stabilità nel tempo. I dati ottenuti hanno evidenziato una degradazione parziale (**Figura 6**, campioni SN5 e SN7) o totale (**Figura 6**, campioni SN2 e SN4) dopo 12 mesi (**Figura 6**).

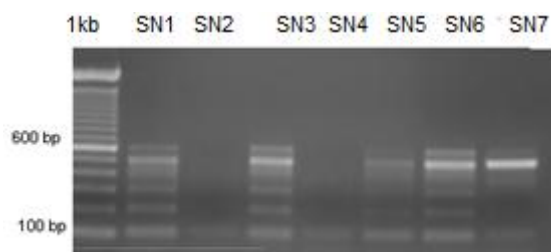


Figura 6. Gel di agarosio 1% degli ampliconi di 7 campioni di DNA estratti da cellule della bocca, ottenuti mediante PCR multiplex (DNA OK, Labogen), dopo 12 mesi.

4.1.3. Utilizzo DNA estratto da sangue vs DNA estratto da cellule della bocca.

Per valutare se il DNA da cellule della bocca è utilizzabile in biologia molecolare, 10 campioni di DNA estratto da sangue e da cellule della bocca provenienti dallo stesso individuo sono stati sottoposti a:

- ✓ Sequenziamento diretto per il polimorfismo del gene *ABCC8*
- ✓ Saggi di discriminazione allelica per i polimorfismi dei geni *GSTP1* e *CASR*

Come riportato in **Tabella 8**, dall'analisi di sequenziamento diretto e di discriminazione allelica dei suddetti campioni, è stata riscontrata una concordanza del 100% tra i genotipi delle due classi di campioni esaminate.

	<i>Sangue (n=10)</i>			<i>Cellule della bocca (n=10)</i>		
Sequenziamento diretto	G/G			G/G		
SNP rs17846762 (gene <i>ABCC8</i>)	10			10		
Saggio TaqMan	T/G	T/T	G/G	T/G	T/T	G/G
SNP rs46078013 (gene <i>CASR</i>)	5	0	5	5	0	5
Saggio TaqMan	A/G	A/A	G/G	A/G	A/A	G/G
SNP rs1695 (gene <i>GSTP1</i>)	3	7	0	3	7	0

Tabella 8. Dati di genotipizzazione ottenuti mediante sequenziamento diretto e saggi TaqMan su 10 campioni di DNA estratti da sangue e da cellule della bocca.

4.2. Valutazione dell'utilizzo e affidabilità della WGA in biologia molecolare su campioni di DNA da sangue e da cellule della bocca

4.2.1. Confronto tra due kit commerciali

Dopo WGA la quantità di DNA ottenuta da 22 ng di DNA genomico varia da 17,6 µg a 36,6 µg utilizzando il kit *Repli-g Mini* (Qiagen) e da 32,6 µg a 60,1 µg utilizzando il kit *GenomiPhi* (GE Healthcare) (**Tabella 9**).

	Kit Repli-g Mini (Qiagen)	Kit GenomiPhi (GE-Healthcare)
Campioni	Resa DNA (µg)	Resa DNA (µg)
N1G	17,6	32,6
N2G	20,9	52,8
N3G	17,7	34,1
N4G	23,4	36,5
N5G	27,8	42,6
N6G	24,8	39,4
N7G	26,8	48,8
N8G	32,4	60,1
N9G	36,6	40,6
N10G	32,0	35,6

Tabella 9. Rese di DNA (µg) di 10 campioni di DNA da sangue amplificati con il kit *Repli g-Mini* e il kit *GenomiPhi*.

Con il kit *Repli-g* è stata ottenuta un'amplificazione di 230 volte, mentre con il kit *GenomiPhi* l'amplificazione è stata di 1500; i fattori di amplificazione di ciascun campione ottenuti con entrambi i kit sono riportati in **Figura 7**.

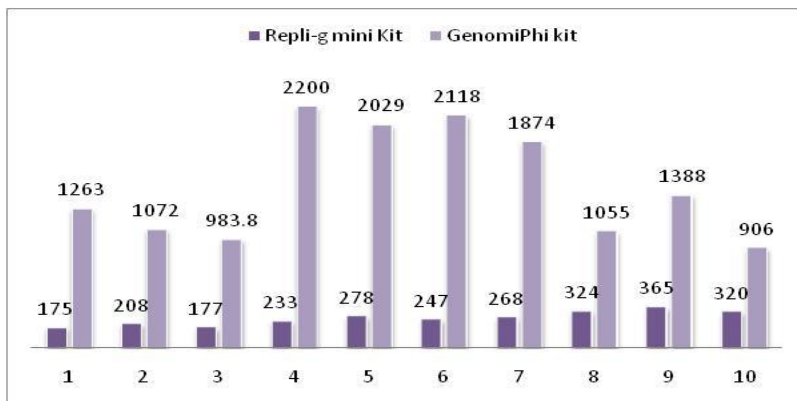


Figura 7. Fattori di amplificazione ottenuti da 10 campioni di DNA estratti da sangue amplificati con i kit *Repli-g Mini Kit* (Qiagen) e *GenomiPhi kit* (GE Healthcare).

4.2.2. Purificazione dei prodotti amplificati con WGA.

Le quantità di prodotto purificato sono state misurate al Nanodrop ND-100. La resa media di DNA dopo purificazione con sodio/acetato è stata di 0.82 µg, mentre la resa media dopo purificazione mediante colonnine è stata di 9 µg. Le rese dei singoli campioni sono riportate in **Tabella 10**.

<i>Campione</i>	<i>Resa G-DNA</i>	<i>Resa GP-DNA dopo purificazione sodio/acetato</i>	<i>Resa GP-DNA dopo purificazione colonnine</i>
<i>N1G</i>	17,6 µg	1.3 µg	8.2 µg
<i>N2G</i>	20,9 µg	0.7 µg	10.3 µg
<i>N3G</i>	17,7 µg	0.9 µg	6.5 µg
<i>N4G</i>	23,4 µg	0.8 µg	11.8 µg
<i>N5G</i>	27,8 µg	0.4 µg	8.2 µg

Tabella 10 .Resa (µg) di DNA di campioni purificati mediante purificazione con sodio acetato e mediante colonnine.

4.2.3. Valutazione della qualità e quantità dei prodotti WGA.

Mediante valutazione con PCR “multiplex”, gel di agarosio e lettura spettrofotometrica abbiamo osservato che il DNA dei campioni oggetto di questo studio, amplificati con il kit *GenomiPhi* e purificati mediante colonnine, è stato di ottima qualità (valore medio: A260/280 ~1,8) e quantità (valore medio: 12 µg) (dati non riportati).

4.2.4. Valutazione della fedeltà della WGA

Sequenziamento diretto

Nei campioni di DNA estratti da sangue e da cellule della bocca sono stati genotipizzati i seguenti SNPs:

rs17846762 (G>A), nell'esone 25 del gene *ABCC8*

rs1800629 (G>A) e rs361525 (G>A) nell'esone 2 del gene *TNF- α*

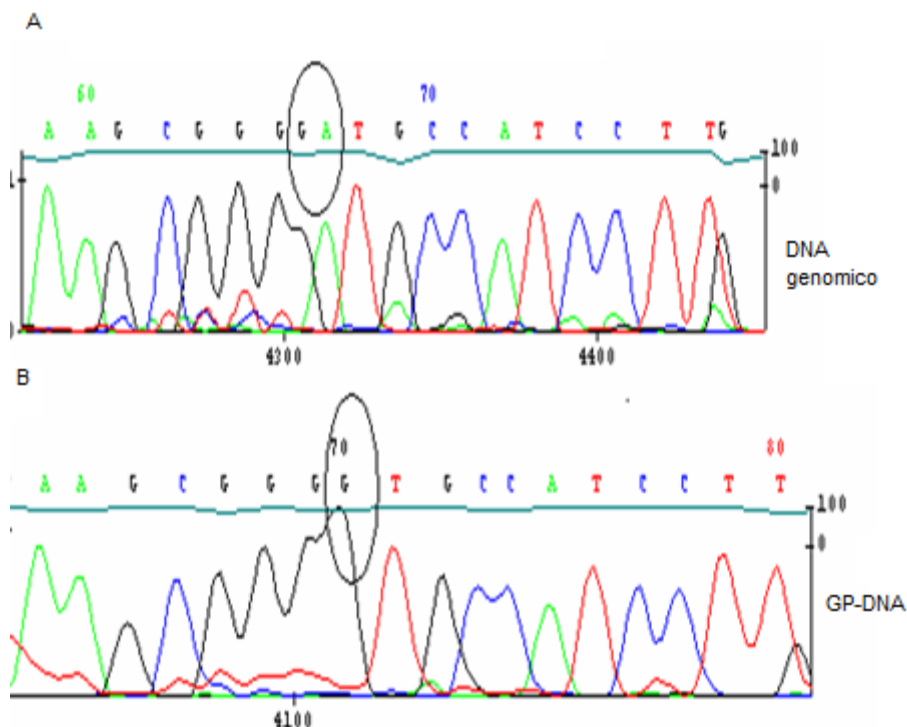
rs1695 (A>G) nell'esone 5 del gene *GSTP1*

L'analisi dei polimorfismi dei geni *ABCC8* e *TNF- α* non ha evidenziato alcuna variazione di genotipo dopo WGA, invece l'analisi del polimorfismo del gene *GSTP1* ha evidenziato una perdita di eterozigotità (A/G>G/G) in 3 campioni di DNA estratti da cellule della bocca su 15 campioni analizzati (**Tabella 11**).

	Prima WGA			Dopo WGA		
	Genotipi	Sangue	Cellule della bocca	Genotipi	Sangue	Cellule della bocca
rs17846762	G/G	10	10	G/G	10	10
	G/G	10	10	G/G	10	10
rs1800629	G/G	12	7	G/G	12	7
	A/G	10	2	A/G	10	2
	A/A	4	0	A/A	4	0
rs361525	G/G	22	9	G/G	22	9
	A/G	4	0	A/G	4	0
	A/A	0	0	A/A	0	0
rs1695	G/G	2	2	G/G	2	5
	A/G	4	9	A/G	4	6
	A/A	10	4	A/A	10	4

Tabella 11. Genotipi ottenuti mediante sequenziamento diretto degli SNP rs1784762, rs1800629, rs361525 e rs1695 in campioni di DNA prima e dopo amplificazione mediante WGA (in rosso sono riportati i dati discordanti.)

Dall'analisi di sequenziamento diretto degli esoni 10, 37 e 39 del gene *ABCC8* in cui era stata riscontrata la presenza di una mutazione senso, in un caso sui tre analizzati si è verificata la perdita di eterozigotità dopo *WGA* (G/A>G/G). In figura 8 è riportato l'elettroferogramma di una regione dell'esone 39 del gene *ABCC8* del suddetto campione prima della *WGA* (eterozigote G>A) e dopo *WGA* (omozigote G>G) (**Figura 8**).



*Figura 8. Elettroferogramma di una regione dell'esone 39 del gene *ABCC8*, A: DNA genomico con mutazione in eterozigosi, B: GP-DNA con perdita di eterozigotità.*

Analisi del microsatellite HUM-FIBRA (FGA).

Dall'analisi del marcatore tetranucleotidico HUM-FIBRA (FGA) è stata osservata una variazione della lunghezza del frammento (bp) dopo amplificazione mediante WGA in 4 campioni sui 16 analizzati (**Tabella 12**).

Campione	PRIMA-WGA		DOPO WGA	
	allele1 (bp)	allele2 (bp)	allele1 (bp)	allele2 (bp)
VM281161	184	201	188	192
FL131163	188	192	188	193
ZL280943	196	196	196	196
GM210453	184	196	184	196
CF151253	204	210	204	210
GO080945	184	188	184	188
ML291262	180	196	188	205
MG310359	200	204	196	201
LR280466	200	204	200	204
VM080371	188	192	188	192
CM120749	200	204	200	204
AG121127	192	196	192	196
LA010449	184	192	184	192
OE101159	192	200	192	200
GA210646	184	188	184	188
PP090848	188	205	180	197

Tabella 12. Analisi del microsatellite FGA del gene del fibrinogeno nei campioni di DNA estratti da sangue prima e dopo amplificazione con WGA (i campioni che presentano una variazione sono evidenziati in rosso).

Saggi di discriminazione allelica mediante “TaqMan Assay” (Applied Biosystems).

Dall'analisi di discriminazione allelica per i due polimorfismi dei geni *GSTP1* e *CaSR* non abbiamo riscontrato nessuna variazione di genotipo prima e dopo amplificazione mediante WGA (**Tabella 13**).

PRE-WGA					POST- WGA										
		Sangue (n=26)			Cellule della bocca (n=10)					Sangue (n=26)			Cellule della bocca (n=10)		
<u>Gene</u>	<u>Genotipi</u>	G/G	17	5	<u>Genotipi</u>	G/G	17	5		<u>Genotipi</u>	G/G	17	5		
<u>CASR</u>		T/T	1	0		T/T	1	0			T/T	1	0		
		T/G	8	5		T/G	8	5			T/G	8	5		
<u>Gene</u>	<u>Genotipi</u>	G/G	2	0	<u>Genotipi</u>	G/G	2	0		<u>Gene</u>	<u>Genotipi</u>	G/G	2	0	
<u>GSTP1</u>		A/A	17	6		A/A	17	6		<u>GSTP1</u>	<u>Genotipi</u>	A/A	17	6	
		A/G	7	4		A/G	7	4		<u>GSTP1</u>	<u>Genotipi</u>	A/G	7	4	

Tabella 13. Analisi dei genotipi mediante saggi di discriminazione allelica su campioni di DNA estratti da sangue e da cellule della bocca prima e dopo amplificazione mediante WGA.

Genotipizzazione mediante Gene Chip Human Mapping 250K Nspl Array.

Nella maggior parte dei campioni di DNA analizzati prima dell'amplificazione la rivelazione degli SNPs (SNPs “call rate”) è stata maggiore del 93% e in alcuni casi superiore al 98%. Ciò indica una buona qualità del saggio.

La “call rate” media dei campioni di DNA genomico è stata 97%, mentre la “call rate” media dei corrispondenti campioni amplificati con WGA è stata 83%. In figura 9 sono riportate le “call rate” di ciascun campione prima e dopo amplificazione con WGA (**Figura 9**).

Dei 262,264 SNPs testati con la piattaforma *Gene Chip Human Mapping 250K Nspl Array*, 38459 (15%) mantengono la concordanza prima e dopo WGA su tutti i campioni esaminati (HI15M *versus* HI15MGP, HI25M *versus* HI25MGP, HI33F *versus* HI33FGP, N5 *versus* N5GP, N5 *versus* SN5GP), mentre 23052 sono sempre discordanti in tutti i campioni prima e dopo WGA. Per i restanti SNP la concordanza varia nelle coppie di campioni analizzati come riportato in Tabella 14.

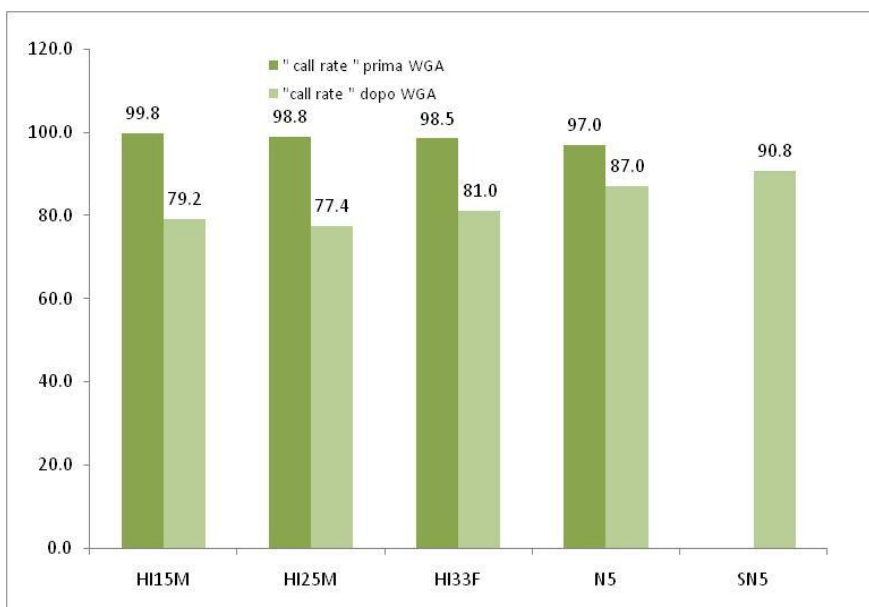


Figura 9. "Call rate" di 4 campioni di DNA da sangue e di un campione da cellule della bocca prima e dopo amplificazione con WGA.

In alcuni campioni è stata riscontrata una diminuzione ed in altri un aumento delle chiamate genotipiche complessive AA, BB, AB, e un sostanziale aumento delle "no call" complessive dopo WGA (**Tabella 14**, pannello: A, B, C, D, E). Ad esempio nel campione HI15M le chiamate AA da 94578 sono diventate 76278 dopo WGA, le chiamate AB da 75211 sono diventate 70307 dopo WGA, le chiamate BB da 91709 sono diventate 61167 dopo WGA mentre le "no call" da 766 sono diventate 54512 dopo WGA.

HI15M							A
HI15MGP		AA	AB	BB	NoCall	Totale complessivo	
	AA	43323	23247	8694	19314	94578	
	AB	20507	22489	16322	15893	75211	
	BB	12210	24361	35991	19147	91709	
	NoCall	238	210	160	158	766	
	Totale complessivo	76278	70307	61167	54512	262264	
	% concordanza	38%					

HI25M							B
HI25MGP		AA	AB	BB	NoCall	Totale complessivo	
	AA	50526	19751	5269	19812	95358	
	AB	14079	29778	9596	18561	72014	
	BB	8051	22465	40944	20256	91716	
	NoCall	750	1051	647	728	3176	
	Totale complessivo	73406	73045	56456	59357	262264	
	% concordanza	46%					

HI33F							C
HI33FGP		AA	AB	BB	NoCall	Totale complessivo	
	AA	61978	16780	3353	16386	98497	
	AB	11627	29094	8083	15010	63814	
	BB	5402	20303	51875	17211	94791	
	NoCall	1262	1743	964	1193	5162	
	Totale complessivo	80269	67920	64275	49800	262264	
	% concordanza	55%					

N5							D
N5GP		AA	AB	BB	NoCall	Totale complessivo	
	AA	70997	12188	113	11793	95091	
	AB	3641	50746	3372	11903	69662	
	BB	161	11622	68168	12063	92014	
	NoCall	706	2823	742	1226	5497	
	Totale complessivo	75505	77379	72395	36985	262264	
% concordanza		72%					

N5							E
SN5GP		AA	AB	BB	NoCall	Totale complessivo	
	AA	85927	3713	75	5376	95091	
	AB	8262	41838	7827	11735	69662	
	BB	131	3446	82341	6096	92014	
	NoCall	1183	1995	1322	997	5497	
	Totale complessivo	95503	50992	91565	24204	262264	
% concordanza		8 %					

Tabella 14. tabelle di contingenza delle chiamate genotipiche dei campioni di DNA genomico e del suo corrispettivo amplificato con WGA (in rosso sono evidenziate le chiamate genotipiche AA, BB, AB e “no call” concordanti nei campioni prima e dopo WGA).

Analisi del numero di Copie (CN) e perdita di eterozigosità (LOH)

Su tutti i campioni GP-DNA (HI15MGP, HI25MGP, HI33FGP) sono state riscontrate numerose amplificazioni (dati non riportati). Tali amplificazioni, non presenti nei campioni di DNA genomico (HI15M, HI25M, HI33F) sono artefatti dovuti alla metodica di amplificazione (dati non riportati). Per esempio, in figura 10 è riportata la mappa genomica del campione HI33F prima e dopo amplificazione WGA, ottenuta con CNAG v3.0, dove è visibile un aumento del numero di copie su tutta la lunghezza del cromosoma (**Figura 10**, pannello in basso: Linea rosa).

A livello di perdita di eterozigosità (LOH), solo il campione di DNA genomico HI33F presenta lunghi tratti di omozigosità (>1 Mb) sui cromosomi 2, 3, 4, 5, 11 e 15 (**Figura 10**). Tali tratti di omozigosità sono stati persi dopo amplificazione con WGA (**Figura 10**). La perdita di questi tratti di omozigosità spiega come mai le chiamate eterozigoti del campione HI33F sono aumentate dopo WGA a scapito di quelle omozigoti che sono diminuite (pannello C, **Tabella 14**). In particolare sul cromosoma 11 due tratti di LOH di 14,5 Mb e 30 Mb sono stati persi dopo amplificazione con WGA (HI33FGP) (**Figura 11**).

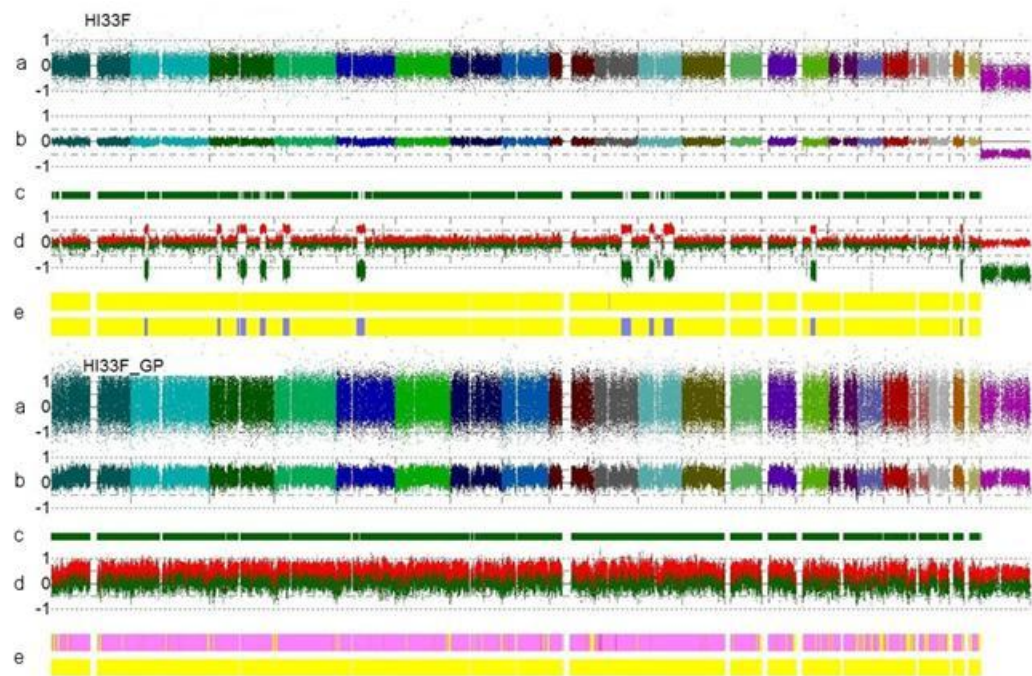


Figura 10. Visualizzazione del campione HI33F (pannello in alto) e del campione HI33FGP (pannello in basso). (a) Ogni cromosoma è rappresentato in differenti colori; (b) la barra rappresenta il segnale di "copy number" di ogni singolo SNP su scala logaritmica, (c) la linea verde rappresenta le chiamate genotipiche eterozigoti, (d) la barra rappresenta il dosaggio allelico dove la linea rossa rappresenta l'allele maggiore e la linea verde l'allele minore, (e). le due barre rappresentano la visualizzazione mediante un codice colore dello stato di copy number (linea gialla: diploidia, linea rosa: acquisto, linea blu: perdita) e della perdita di eterozigotità (LOH) (linea blu: LOH, linea gialla: conservazione della eterozigotità.)

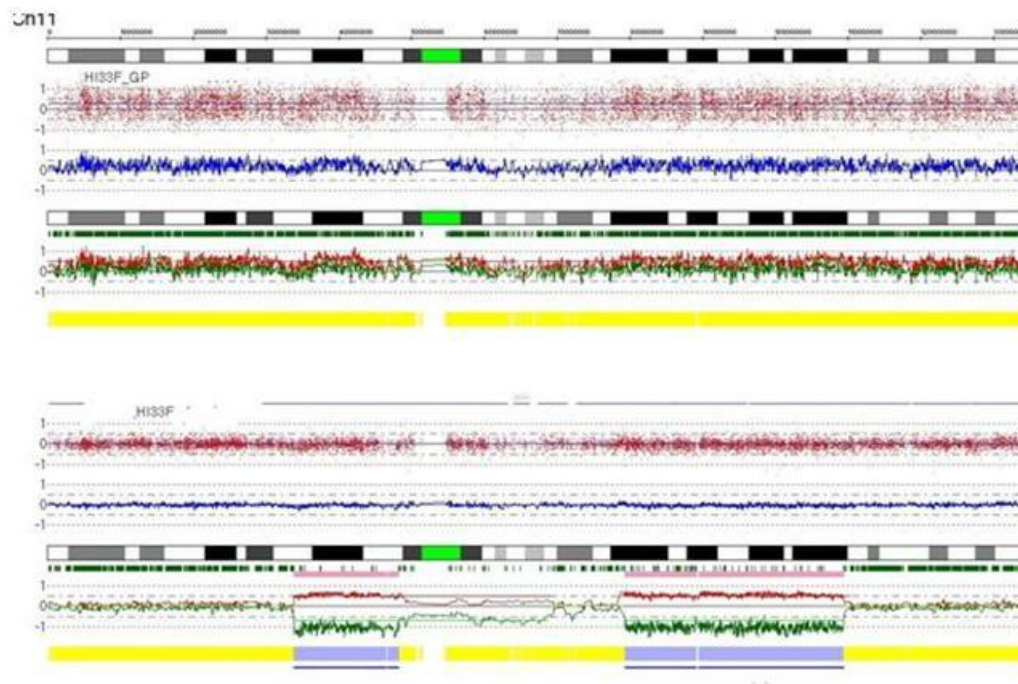


Figura 11. Visualizzazione del chr 11 nei campioni HI33FGP (pannello in alto) e HI33F (pannello in basso) mediante il software CNAG 3.0. Il cromosoma è visualizzato dal braccio p al braccio q (da sinistra verso destra). Le due barre in alto rappresentano il segnale di "copy number" per ogni singolo SNP su scala logaritmica (puntini rossi), e il valore di "Copy number" calcolato su 10 SNP adiacenti (linea blu). La barra verde nel centro rappresenta le chiamate genotipiche eterozigoti e la linea rosa le chiamate omozigoti. Al di sotto è visualizzato il dosaggio allelico dove la linea rossa rappresenta l'allele maggiore e la linea verde l'allele minore. La barra in basso rappresenta una visualizzazione con un codice colore dello stato di LOH (linea gialla: conservazione eterozigotà, linea blu: LOH).

4.3. Studio epidemiologico

L'analisi dei genotipi del polimorfismo rs1695 del gene *GSTP1* su 103 campioni ha evidenziato che 37 campioni (36%) sono omozigoti per l'allele A/A (Ile105/Ile105), 60 campioni (59%) sono eterozigoti per l'allele A/G (Ile105/Val105), e 5 (4.9%) sono omozigoti per l'allele G/G (Val105/Val105) (**Tabella 15**). Le frequenze genotipiche sono state confrontate con quelle di una popolazione toscana reclutata per il progetto HapMap (<http://hapmap.ncbi.nlm.nih.gov/>). L'analisi statistica del χ^2 non ha evidenziato differenze statisticamente significative nella distribuzione dei genotipi del polimorfismo rs1695 del gene *GSTP1* tra i nostri campioni e la popolazione di riferimento considerata ($\chi^2 = 3,92$; g.d.l = 2; p = 0,140).

	Popolazione milanese (n=103)	Popolazione toscana (n=114)
<i>GSTP1</i> Ile105/Ile105 (A/A)	37	53
<i>GSTP1</i> Ile105/Val105 (A/G)	60	52
<i>GSTP1</i> Val105/Val105 (G/G)	5	9

Tabella 15. Rappresentazione dei genotipi nei campioni di cellule della bocca e nella popolazione toscana di riferimento.

5. DISCUSSIONE

Lo scopo principale del mio progetto di dottorato è stato quello di valutare se la *Whole Genome Amplification* potesse essere utilizzata in tutti quei casi in cui si ha una scarsa quantità di DNA sia perché il materiale a disposizione è scarso o potrebbe esaurirsi, sia perché i donatori sono ad esempio bambini cui non possono essere prelevate grandi quantità di sangue, specie se neonati. Ciò ha reso necessaria un'attenta valutazione, mediante controllo di qualità e quantità, dei campioni biologici, in particolare le cellule della bocca, e della WGA con metodiche di biologia molecolare (**Figura 12**). Il sangue è stato sempre il principale fluido biologico da cui ottenere materiale genetico proprio perché il campione di sangue è in grado di fornire una grande quantità di cellule da cui ottenere DNA [38]. Considerando che gli studi genetici possono coinvolgere partecipanti sparsi in diversi paesi e che i donatori possono essere persone anziane e bambini ai quali diventa difficile, se non impossibile prelevare il sangue, il nostro primo obiettivo è stato quello di mettere a punto un metodo di raccolta di materiale biologico alternativo al prelievo di sangue venoso. Di fronte alla possibilità di raccogliere la saliva o le cellule della bocca noi abbiamo scelto di utilizzare le cellule della bocca poiché il nostro obiettivo era di impiegare un metodo di raccolta che consentisse un facile trasporto e conservazione del campione, in letteratura sono riportati pochi dati concernenti l'utilizzo delle cellule della bocca, al contrario ci sono moltissimi articoli in cui è impiegata la saliva in studi epidemiologici [40]. Abbiamo scelto di raccogliere i campioni con le FTA cards, apposite cartine di carta da filtro sulle quali applicare il campione prelevato, spedibili tramite posta senza timore di dispersione di liquidi e conservate in modo idoneo a temperatura ambiente [39]. A tutti i soggetti reclutati, abbiamo fornito il materiale necessario per il prelievo delle cellule della bocca: FTA card, tampone sterile, guanti monouso e relative istruzioni per la procedura di prelievo, deposito del campione sulle cartine e loro conservazione fino al momento della consegna. Al fine di ridurre al minimo la contaminazione batterica abbiamo chiesto ai donatori di prelevare il campione a digiuno, come riportato in letteratura [13]. Dopo avere ottenuto le cartine dai donatori, abbiamo dovuto mettere a punto una metodica di estrazione del DNA. Una prima procedura per l'estrazione del DNA da cellule della bocca è stata quella indicata dalla ditta produttrice delle cartine, ma la quantità di DNA ottenuta è stata pressoché nulla (dati non riportati). Abbiamo allora utilizzato la procedura per l'estrazione del DNA da spot di sangue come indicato dal kit di estrazione in uso presso il nostro laboratorio (Qiamp DNA Mini kit Qiagen). In accordo con la letteratura [19,40], il DNA che abbiamo ottenuto dai campioni di cellule della bocca è stato di scarsa quantità e qualità. La scarsa qualità del DNA ottenuto da cellule della bocca è dovuta a contaminazione con fenoli, carboidrati e altri composti aromatici derivanti dal metodo di raccolta utilizzato. Inoltre non è stato possibile visualizzare DNA ad alto peso molecolare nonostante la lettura spettrofotometrica

abbia indicato la presenza di una quantità di DNA variabile (8-28 ng/μl), costituito sia da DNA umano sia da DNA batterico. La presenza di DNA batterico è stata dimostrata mediante amplificazione del gene che codifica l'RNA batterico (16s) e rilevando la presenza di tale amplificato nei campioni di cellule della bocca (dati non riportati). Dawes C [41] ha riportato che, sebbene la percentuale di DNA batterico sia consistente (la saliva contiene $1,7 \times 10^7$ /ml di cellule batteriche), questo non ha inficiato la riuscita delle successive analisi di biologia molecolare. Lo stesso risultato è stato confermato da Dominique Quinque e altri autori che hanno quantificato mediante PCR quantitativa il DNA batterico presente nei campioni di cellule della bocca per valutare l'eventuale interferenza [42]. Nel nostro studio non è stato necessario quantificare il DNA batterico mediante PCR quantitativa, poiché siamo riusciti a utilizzare tali campioni per reazioni di amplificazione e per analisi di genotipizzazione. In accordo con i risultati ottenuti da Walker [43], abbiamo riscontrato una degradazione del DNA nell'arco di 12 mesi, causata dalla presenza di proteine, carboidrati e fenoli; contrariamente dopo 6 mesi Lum [44] non ha osservato degradazione del DNA. Per gli studi a lungo termine sarà importante valutare se il DNA estratto si degrada nell'arco di un certo tempo e se utilizzando metodi di estrazione diversi si possa evitare questo inconveniente. Recentemente Helene Johanson [45] e altri autori hanno riportato che la fissazione della cartina con metanolo non solo impedisce la degradazione del DNA nel tempo, ma ne migliora la quantità e qualità ottimizzando anche, in termini di costi, l'uso delle FTA cards per la raccolta di campioni. La fissazione delle cartine con metanolo, da noi valutata su alcuni campioni, non ha prodotto DNA di migliore qualità. Resta comunque da valutare se il DNA estratto da cartine fissate con metanolo presenta una maggiore stabilità nel tempo.

Alla luce dei vantaggi e svantaggi legati all'utilizzo del DNA estratto dalle cellule della bocca, noi abbiamo voluto valutare se il DNA potesse essere impiegato in biologia molecolare per analisi di sequenziamento diretto, saggi di discriminazione allelica e analisi "genome wide" con gli "SNP Array", comparabilmente al DNA estratto da sangue. A tale scopo abbiamo reclutato 10 soggetti sani che hanno donato sia un campione di sangue sia un campione di cellula della bocca raccolto tramite FTA card per effettuare analisi in parallelo. I nostri risultati hanno dimostrato che tutti i campioni di DNA estratti dalle cellule della bocca sono amplificabili mediante PCR, nonostante siano di quantità e qualità inferiore rispetto ai campioni di DNA da sangue; a differenza da quanto riportato da Hansen e Coll [19] che non sono riusciti ad amplificare nessuno dei campioni di DNA ottenuti da FTA card. Inoltre il 100% dei nostri campioni è stato genotipizzato mediante saggio TaqMan e sequenziamento diretto, mentre Hansen e Coll hanno ottenuto la genotipizzazione solo nel 23% dei campioni di DNA ottenuti da FTA card. Non abbiamo invece potuto utilizzare il DNA dalle cellule della bocca per l'analisi "genome wide" con gli "SNP Array" poiché la metodica da noi utilizzata (Affymetrix Gene Chip Human Mapping 250K) richiede un DNA di

ottima qualità non ottenibile con la modalità di prelievo da noi utilizzata (Tabella 7). Al contrario, Andrew E ed altri autori [46], hanno dimostrato che il DNA estratto dalle cellule della bocca è utilizzabile al pari del DNA estratto da sangue per l'analisi di "copy number" utilizzando la piattaforma Illumina. Per quanto riguarda i risultati di genotipizzazione mediante "TaqMan Assay" e sequenziamento diretto, abbiamo ottenuto il 100% di concordanza dei genotipi tra DNA da cellule della bocca e campioni di DNA da sangue confermando i risultati ottenuti da altri autori [13]. Concludendo, le cellule della bocca rappresentano un campione biologico da impiegare in studi epidemiologici in alternativa al campione di sangue. Inoltre il DNA estratto dalle cellule della bocca in alcuni contesti, è compatibile con applicazioni "genome wide" che stanno diventando sempre più di routine in studi di popolazione o di associazione [42]. Tuttavia la quantità di materiale ottenuta da tale risorsa biologica rimane scarsa e limita il numero di analisi fattibili sul campione, da qui la necessità di utilizzare un metodo per incrementare la quantità di DNA, quale la "Whole Genome Amplification" (WGA) che si basa su un processo di "Amplificazione mediante Spostamento Multiplo del filamento" (MDA) [49,50].

Tra i diversi metodi di WGA noi abbiamo scelto di utilizzare un metodo WGA non basato su una reazione di PCR in modo tale che assicurasse una bassa percentuale di errore rispetto ai metodi WGA basati su una reazione di PCR come la PEP e la DOP-PCR che introducono un'elevata percentuale di artefatti. Le applicazioni possono essere molteplici: scienza forense, diagnosi prenatale, "immortalizzazione" di campioni clinici, studi di genotipizzazione, tutti quei casi in cui la quantità del DNA è scarsa già in partenza o si esaurisce nel corso delle analisi [47]. Tuttavia la WGA deve soddisfare alcuni requisiti di base: l'amplificazione dev'essere molto accurata in modo da evitare errori, il processo di amplificazione non deve produrre un errore nella distribuzione dei prodotti di DNA, infine è necessario ottenere un elevato fattore di amplificazione in modo che la WGA produca un'adeguata quantità di DNA partendo da una minima quantità di DNA iniziale [6, 48]. La questione chiave circa l'uso della WGA è questa: è possibile ottenere dopo WGA un prodotto che sia un'accurata riproduzione della sequenza genomica originale? La WGA introduce degli artefatti?

In questo studio abbiamo valutato l'affidabilità della WGA su DNA ottenuto da campioni di sangue e cellule della bocca.

In commercio esistono due kit di WGA basati sul processo di MDA: kit *Repli-g Mini* (Qiagen) e il kit *GenomiPhi* (GE Healthcare). Nel nostro studio abbiamo valutato entrambi i kit su 10 campioni di DNA estratti da sangue ed abbiamo riscontrato che, mentre con il primo kit si ottiene un'amplificazione di 260 volte, con il secondo kit il DNA è amplificato 1500 volte e quindi la resa è molto maggiore. Dopo aver scelto il kit commerciale con la resa più alta, abbiamo valutato la fedeltà e riproducibilità della "Whole Genome Amplification". A tale scopo abbiamo scelto alcune metodiche comunemente utilizzate in biologia molecolare quale sequenziamento diretto, genotipizzazione mediante microsatelliti e saggi di

discriminazione allelica, e metodiche più complesse quale analisi “genome wide” con gli “SNP Array”. Le nostre valutazioni sono state eseguite confrontando i campioni di sangue e cellule della bocca prima e dopo WGA. Sebbene la WGA sia da qualche tempo utilizzata in biologia molecolare ed esistano in letteratura molti dati che riportano l'utilizzo dei campioni amplificati con WGA nel sequenziamento diretto, nei saggi di discriminazione allelica e nell'analisi dei microsatelliti ci sono, invece, pochi dati riguardanti il confronto dei campioni di DNA, sia estratti da sangue sia da cellule della bocca, prima e dopo amplificazione con WGA [49, 50]. In accordo con alcuni dati della letteratura [51] abbiamo rilevato che la WGA può determinare la perdita di eterozigotità per sequenziamento diretto. Anche se questo dato è stato da noi verificato solo in pochi casi, tuttavia data la perdita dell'informazione genetica l'uso della WGA nel sequenziamento diretto nel caso di ricerca di mutazioni causative di malattie è sconsigliato. I dati presenti in letteratura relativi alla fedeltà della WGA sono molto contrastanti, alcuni ricercatori ritengono che la WGA possa essere utilizzata nel sequenziamento poiché non determina cambiamenti nella sequenza del DNA, altri ricercatori sostengono che la WGA determini una perdita di eterozigotità nella sequenza di geni molto polimorfici [51]. È opportuno evidenziare che nessuno ha mai utilizzato DNA amplificato mediante MDA per l'analisi di mutazioni di un gene candidato ma l'utilizzo è stato rivolto a studi di tipo epidemiologico i quali sono effettuati su un numero elevato di campioni e possono ammettere un errore trascurabile senza inficiare il risultato. Le discordanze osservate nei campioni amplificati sono presenti sia su campioni di DNA estratto da sangue sia su DNA estratto da cellule della bocca (Tabella 11). Diversamente, nell'analisi di discriminazione allelica su due polimorfismi valutati in due geni differenti (*GSTP1* e *Calcium Sensing Receptor*) abbiamo riscontrato il 100% di concordanza dei genotipi nei campioni prima e dopo amplificazione con WGA, in accordo con i dati riportati in letteratura [51]. In precedenza alcuni autori hanno dimostrato che la genotipizzazione mediante saggio “TaqMan Assay” ha dato il 99,95 % di concordanza di genotipi [7, 52] prima e dopo WGA, tali differenze sono state però attribuibili ad un errore di genotipizzazione e non alla metodica WGA.

Per quanto riguarda la possibilità di utilizzare i campioni amplificati con WGA in analisi di microsatelliti, anche in questo caso i nostri dati hanno dimostrato una discordanza tra i campioni prima e dopo WGA. In particolare noi abbiamo scelto di analizzare il microsatellite HUM-FIBRA (FGA) del gene del fibrinogeno, il quale è un marcatore tetra nucleotidico molto polimorfico. In 4 dei 16 campioni di DNA estratti da sangue, analizzati prima e dopo amplificazione, abbiamo riscontrato una variazione nella lunghezza espressa in paia di basi (bp) dopo amplificazione. Secondo Dickson [53] ciò è dovuto al fatto che il processo di WGA può introdurre delle delezioni o espansioni dei microsatelliti attraverso lo slittamento della replicazione mediata dalla polimerasi. Noi ci siamo limitati ad analizzare il microsatellite solo su campioni di DNA estratti da sangue poiché in letteratura

Guangyun Sun [54] ha riportato che nei campioni di cellule della bocca, raccolte mediante le FTA Cards, si crea uno sbilanciamento allelico dovuto ad un'interazione sconosciuta tra i reagenti del kit di amplificazione *GenomiPhi* e le FTA Cards. Considerando che i microsatelliti sono impiegati nei laboratori di analisi per stabilire la paternità, si può concludere che la WGA non può essere impiegata perché altera la veridicità del test.

Valutando la fedeltà della WGA mediante un'analisi "genome wide" impiegando la piattaforma "Affymetrix Gene Chip Human Mapping 250K", abbiamo riscontrato nei nostri campioni analizzati, una diminuzione delle "call rate" dopo WGA; ciò può dipendere da diversi fattori: la metodica WGA basata sulla MDA può introdurre un elevato rumore di fondo rispetto ad altri tipi di metodiche e questo non permette di attribuire chiamate genotipiche reali, inoltre noi abbiamo interrogato tutti i 262264 SNPs, mentre altri autori eliminano gli SNPs che danno chiamate genotipiche discordanti [55]. Un dato rilevante emerso dai nostri studi è che su 262264 SNPs 38459 SNPs concordanti tra i campioni di DNA genomico e GP-DNA per cui se si compie un'analisi circoscritta a questi SNP la WGA assicura una conservazione di genotipo.

Paez e altri autori [56] in uno studio condotto impiegando gli SNP Array 10K hanno riscontrato una concordanza di genotipo tra i campioni di DNA genomico e i campioni amplificati mediante WGA >99%. Questa percentuale di concordanza è stata però ottenuta solo nei campioni sottoposti ad una denaturazione alcalina prima dell'amplificazione mediante WGA, infatti, gli stessi autori hanno trovato differenze tra le chiamate prima e dopo WGA nei campioni non denaturati prima del processo di amplificazione mediante WGA. Jasmine ed altri autori [55] hanno ottenuto una "call rate" del 97% nei campioni amplificati con WGA, impiegando la piattaforma "Affymetrix Gene Chip Human Mapping 250K" genotipizzando solo 224,000 SNPs. Inoltre anche Randi ed altri autori [57] hanno ottenuto una "call rate" del 98% su genotipizzando solo 1744 SNPs su 1920 SNPs utilizzando la piattaforma Illumina. Tuttavia la strategia, utilizzata da alcuni ricercatori, di eliminare gli SNPs che danno chiamate genotipiche discordanti non dà risultati fedeli alla realtà.

A livello di perdita di eterozigotità (LOH) abbiamo riscontrato nel campione HI33F, recante tratti di perdita di eterozigotità (LOH) nei cromosomi 2, 3, 5, 11 e 15, la perdita di tali tratti dopo WGA (Figura 10). Analizzando le chiamate genotipiche di tutti gli SNPs compresi nel tratto di LOH di 14.5 Mb, alcune chiamate omozigoti, dopo WGA, diventavano eterozigoti o "no call".

I nostri campioni di DNA genomico non presentano alcuna variazione di "copy number", per cui le amplificazioni presenti su tutti i cromosomi dei campioni amplificati con WGA sono artefatti dovuti a rumore di fondo causato dalla metodica WGA. I dati riportati in letteratura sono molto discordanti: Jason J nel 2008 [57] ha scoperto in 532 campioni amplificati mediante MDA 438 variazioni di "copy number", non riportate nel database delle variazioni di "copy number" ("Genomic

Variants”) [58]. Altri studi precedenti [55] non hanno evidenziato alcuna variazione di “copy number” confrontando campioni prima e dopo amplificazione mediante MDA. In questi studi, comunque, le analisi sono sempre state fatte “normalizzando” i dati sui campioni non amplificati con WGA. Bisogna rilevare che ciò non corrisponde alla realtà, poichè lo scopo principale della WGA è di amplificare il DNA per ottenerne una maggiore quantità quando questa è scarsa o nel caso in cui quel determinato campione è prezioso per il proseguimento di studi successivi. Il fatto che su 262264 SNP ci sia un numero così elevato di errori introdotti dalla WGA a livello di genotipo, di CN e di LOH denota che gli “SNP Gene Chip Human Mapping 250K Nspl Array” non possono essere impiegati per un’analisi su larga scala, ragionevolmente si può pensare di impiegare *array* a maggiore risoluzione che probabilmente riducono gli errori introdotti dalla WGA, sebbene non ci siano dimostrazioni pubblicate in questa direzione.

In letteratura è riportato che il livello di errore introdotto durante la reazione di WGA può essere dovuto alla quantità di materiale di partenza. Ad esempio i ricercatori, Dean e Lovmar [7], hanno valutato le frequenze dei genotipi in campioni amplificati mediante WGA partendo da diverse concentrazioni di materiale iniziale ed entrambi gli autori hanno concluso che si ottiene una perfetta concordanza di genotipi prima e dopo WGA impiegando circa 3 ng di materiale di partenza. Al contrario Bergen ed altri ricercatori [59] hanno valutato la concordanza di genotipi utilizzando diverse concentrazioni di DNA (10, 25, 100 ng) e hanno concluso che si ottiene il 100% di concordanza impiegando 25 ng di DNA genomico di partenza per la reazione di WGA [40]. Alla luce di questi dati riportati in letteratura, gli errori introdotti dalla WGA nel sequenziamento diretto e nell’analisi di microsatelliti potrebbero essere dovuti alla quantità di DNA iniziale impiegata. Il fatto che la WGA introduca errori solo quando amplifica alcune regioni rispetto ad altre può essere imputato al fatto che in quelle specifiche regioni amplifica in modo aspecifico [60]. Fattori come il contenuto di GC della regione amplificata e la posizione cromosomica del polimorfismo possono influenzare la “performance” della WGA determinando un errore di amplificazione [61, 62].

Dalle nostre analisi è emerso che la WGA può essere una risorsa utile per molti ricercatori che si trovano di fronte a problemi quali la scarsa quantità di DNA o l’esaurimento stesso del campione, per cui può essere impiegata per analisi di discriminazione allelica su un numero elevato di polimorfismi.

Dopo una scrupolosa analisi dei campioni di cellule della bocca mediante controlli di qualità e quantità e una serie di valutazioni, mediante alcune tecniche di biologia molecolare, sull’utilizzo della WGA, come riportato nel diagramma a flusso di figura 12, abbiamo raccolto le cellule della bocca di 103 individui residenti nell’area milanese.

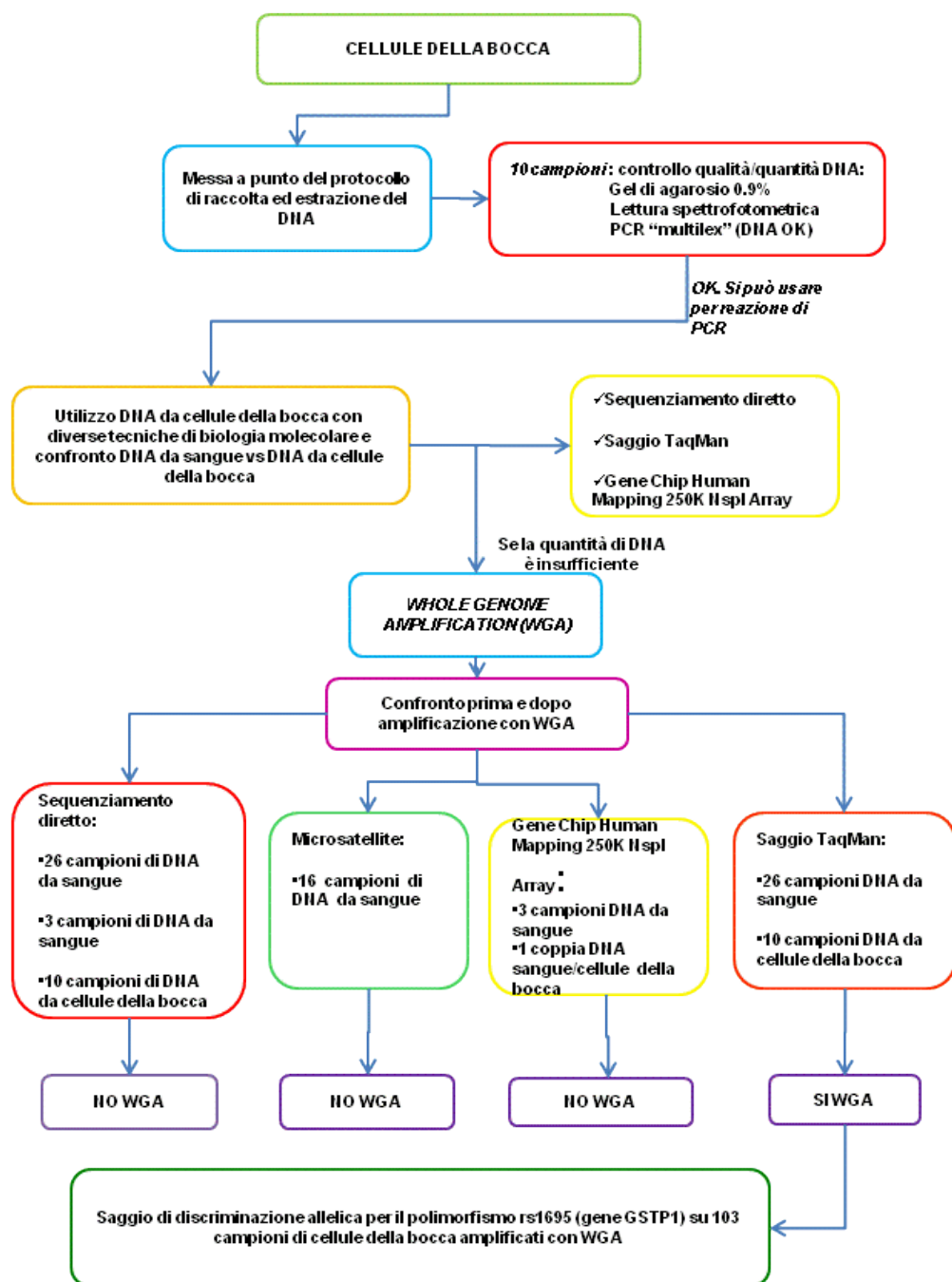


Figura 12. Diagramma di flusso riguardante lo studio svolto

Nell'ambito di un progetto in corso nel nostro laboratorio, intento a studiare le interazioni tra particolari polimorfismi e la predisposizione verso patologie respiratorie, abbiamo amplificato mediante *WGA* i campioni di DNA e studiato il polimorfismo Ile105/Val105 del gene *GSTP1* mediante saggi di discriminazione allelica. Le frequenze dei genotipi della nostra popolazione non sono significativamente differenti da quelli di una popolazione di riferimento toscana reclutata per il progetto *HapMap*.

Poiché non solo i fattori genetici ma anche altri fattori, quali condizioni climatiche, dose e tempo di esposizione all'aria inquinata, età, sesso, influiscono sui meccanismi con cui gli inquinanti presenti nell'aria possono provocare danni alla salute, è di grande utilità avere una popolazione sana della stessa area geografica della popolazione malata pertanto i nostri campioni, quindi, saranno impiegati come controlli sani per gli studi d'interazione gene-ambiente in programma nel nostro laboratorio. Tale studio è svolto su un'ampia casistica d'individui distribuiti su un vasto territorio, e coinvolgerà l'analisi di numerosi SNPs associati all'inquinamento. Da ciò si evidenzia l'importanza della messa a punto della tecnica di estrazione del DNA da cellule della bocca raccolte mediante FTA cards e successiva amplificazione mediante *WGA*.

6. CONCLUSIONI

Dai nostri studi è emerso che le cellule della bocca possono essere usate in alternativa al sangue. Inoltre la raccolta delle cellule della bocca mediante le FTA Cards è stato un metodo molto efficace e non invasivo poichè ci ha permesso di eseguire i nostri studi su un'ampia popolazione dato l'elevato consenso ricevuto, cosa che può essere difficile nel caso di prelievo di sangue. Il DNA estratto dalle cellule della bocca può essere impiegato in analisi genetiche, nonostante la scarsa qualità e quantità e la contaminazione batterica.

In secondo luogo la "Whole Genome Amplification" (WGA) rappresenta un'ottima strategia per risolvere il problema legato alla scarsa quantità di materiale biologico. Tuttavia il suo utilizzo dev'essere preceduto da una valutazione in base all'analisi molecolare da compiere. La compatibilità della WGA con il saggio di discriminazione allelica, dimostrata nel nostro studio, ci ha permesso di dimostrare che l'analisi di genotipizzazione su un numero elevato di SNPs può essere eseguita da materiale biologico di scarsa qualità e quantità.

7. RINGRAZIAMENTI

Al termine del mio percorso, durato tre anni, desidero ringraziare tutte le persone che mi hanno aiutato e mi sono state vicine.

Vorrei ringraziare innanzitutto, la Dr.ssa Cristina Battaglia che è il mio tutore, per la disponibilità a coordinare il mio progetto di dottorato e per il suo contributo nella produzione scientifica.

Un ringraziamento speciale alla Dr.ssa Marica Proverbio per la sua preziosissima competenza in ambito genetico e per il suo aiuto e sostegno in ogni circostanza.

Inoltre ringrazio il mio gruppo di lavoro, in particolare la Dr.ssa Alessandra Gessi e la Dr.ssa Eleonora Mangano che hanno contribuito all'analisi "genome wide".

Ringrazio inoltre Marco Pellegrini per il prezioso aiuto nella formattazione della tesi.

Ringrazio la Dr.ssa Rosanna Asselta e il suo gruppo di lavoro per il loro contributo in alcune analisi senza le quali molti risultati non sarebbero stati ottenuti.

Ringrazio tutti i volontari che ci hanno fornito i campioni di cellule della bocca con entusiasmo di partecipare ad un progetto sulla qualità dell'aria che respiriamo.

Vorrei ringraziare il Prof Scolastico per avermi dato l'opportunità di svolgere questi tre anni di dottorato di ricerca.

La borsa di studio e' stata finanziata dal centro interdipartimentale Studi Biomolecolari e Applicazioni Industriali (CISI) su finanziamento del Comune di Milano (Convenzione 55/2008)

Il lavoro è stato sostenuto dal Comune di Milano impegnato sul fronte della tutela della salute dei cittadini, integrando le politiche della Mobilità, dell'Ambiente, della Ricerca e dell'Innovazione, con uno studio approfondito delle malattie provocate dall'inquinamento (Progetto Prolife).

Oltre ai miei colleghi desidero ringraziare la mia famiglia che mi è sempre stata vicina e mi ha sostenuto sempre.

8. BIBLIOGRAFIA

1. Lovmar L, Syvänen AC." Multiple displacement amplification to create a long-lasting source of DNA for genetic studies", *Hum Mutat*,vol.27,n.7,2006,pp.603-14.
2. Hughes S, S.Arnason N, Done S, Squire J," The use of whole genome amplification in the study of human disease", *Prog Biophys Mol Biol*,vol 88,n.1,2005, pp.173-89.
3. Telenius H, Carter NP, Bebb CE et al," Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer, *Genomics*, vol.133, 1992,pp:718-25
4. Zhang L, Cui X, et al, "Whole genome amplification from a single cell: implications for genetic analysis",*Proc Natl Acad Sci U S A*, vol. 89,1992, pp:5847-51.
5. Lucito R, Nakimura M, West JA et al," Genetic analysis using genomic representations", *Proc Natl Acad Sci U S A*,vol.95, 1998, pp:4487-92.
6. Lasken RS, Egholm M, "Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens", *Trends Biotechnol*,vol.21,n.12, ,2003. pp.531-5.
7. Dean FB, Hosono S, Fang L et al," Comprehensive human genome amplification using multiple displacement amplification", *Proc Natl Acad Sci U S A*,vol.99,n.8,2002,pp.5261-6.
8. Iscove NN, Barbara M et al, "Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA", *Nat Biotechnol*, vol.20, 2002, pp:940-3.
9. Paul M. Lizardi et al," Mutation detection and single-molecule counting using isothermal rolling-circle amplification", *Nature Genetics*,vol.19, 1998,pp.225 – 232.
10. Lasken RS," Genomic DNA amplification by the multiple displacement amplification (MDA) method",*Biochem Soc Trans*,vol. 37,2009,pp. 450-3.
11. Detter JC, Jett JM, Lucas SM," Isothermal strand-displacement amplification applications for high-throughput genomics", *Genomics*, vol.80,2002, 691-8.
12. Hosono S, Faruqi AF, Dean FB et al," Unbiased whole-genome amplification directly from clinical samples", *Genome Res*, vol.13,2003, pp:954-64.
13. Harty LC, Garcia-Closas M, Rothman N et al," Collection of buccal cell DNA using treated cards.",*Cancer Epidemiol Biomarkers Prev.*,vol.9,n.5,2000,pp.501-6.
14. Elizabeth Milne,1 Frank M. van Bockxmeer et al," Buccal DNA Collection: Comparison of Buccal Swabs with FTA Cards", *Cancer Epidemiol Biomarkers Prev*,vol 15,n.4,2006.

15. García-Closas M et al, "Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash", *Cancer Epidemiol Biomarkers Prev*, vol.10,n.6,2001,pp.687-96.
16. Beckett SM, Laughton SJ, Pozza LD et al, "Buccal swabs and treated cards: methodological considerations for molecular epidemiologic studies examining pediatric populations", *Am J Epidemiol*, vol.167,n.10,2008,pp.1260-7.
17. Van Wieren-de Wijer DB, et al, "Determinants of DNA yield and purity collected with buccal cell samples.", *Eur J Epidemiol*, vol.24,n.11,2009,pp.677-82.
18. Feigelson HS, Rodriguez C, "Determinants of DNA yield and quality from buccal cell samples collected with mouthwash", *Cancer Epidemiol Biomarkers Prev*, vol. 10,n.9,2001,pp.1005-8.
19. Hansen TV, Simonsen MK, et al, "Collection of blood, saliva, and buccal cell samples in a pilot study on the Danish nurse cohort: comparison of the response rate and quality of genomic DNA.", *Cancer Epidemiol Biomarkers Prev*, vol.16,n.10,2007,pp.2072-6.
20. Ruzza P, Rosato A, Rossi CR, Floreani M, Quintieri L, "Glutathione transferases as targets for cancer therapy.", *Anticancer Agents Med Chem*, vol.9,n.7,2009,pp.763-77.
21. Lo HW, Ali-Osman F, "Genetic polymorphism and function of glutathione S-transferases in tumor drug –resistance.", *Curr Opin Pharmacol*, vol.7,n.4,2007,pp.367-74.
22. Di Paolo OA, Teitel CH, et al, "Expression of cytochromes P450 and glutathione S-transferases in human prostate, and the potential for activation of heterocyclic amine carcinogens via acetyl-coA-, PAPS- and ATP-dependent pathways", *Int J Cancer*, vol.117, 2005, pp:8-13.
23. McIlwain CC, Townsend DM, Tew KD, "Glutathione S-transferase polymorphisms: cancer incidence and therapy", *Oncogene*, vol.25,2006, pp:1639-48.
24. Lo HW, Ali-Osman F, "Genetic polymorphism and function of glutathione S-transferases in tumor drug resistance", *Curr Opin Pharmacol*, vol.7, 2007, pp:367-74.
25. Del-Rio-Navarro B, Berber A, Blandón-Vijil V, Ramírez-Aguilar M, Romieu I et al, "Identification of asthma risk factors in Mexico City in an International Study of Asthma and Allergy in Childhood survey.", *Allergy Asthma Proc*, vol. 27,n.4,2006,pp.325-33.
26. Yang IA, Fong KM et al, "Genetic susceptibility to the respiratory effects of air pollution", *Thorax*, vol.63, 2008, pp:555-63.
27. Piacentini S, Polimanti R, Moscatelli B et al, "Glutathione S-transferase gene polymorphisms and air pollution as interactive risk factors for asthma in a multicentre Italian field study: A preliminary study", *Ann Hum Biol*, VOL.37, 2010, PP:427-39.

28. Gilliland FD, Li YF, Saxon A, Diaz-Sanchez D," Effect of glutathione-S-transferase M1 and P1 genotypes on xenobiotic enhancement f allergic responses: randomised, placebo-controlled crossover study.",*Lancet*10,vol.363,n.9403,2004,pp.119-25.
29. Millard TP, Fryer AA, et al," A protective effect of glutathione-S-transferase GSTP1*Val(105) against polymorphic light eruption", *J Invest Dermatol*, vol.128, 2008, 2008, pp:1901-5.
30. Tamer L, Calikoğlu et al, "Glutathione-S-transferase gene polymorphisms (GSTT1, GSTM1, GSTP1) as increased risk factors for asthma", *Respirology*, vol.9, 2009, pp:493-8.
31. Kleeberger SR.," Genetic aspects of susceptibility to air pollution.",*Eur Respir J Suppl*,vol.40,2003,pp.52s-56s
32. Churg A.," Interactions of exogenous or evoked agents and particles: the role of reactive oxygen species.", *Radic Biol Med*,vol.34,n.10,2003,pp.1230-5.
33. Kleeberger SR, Levitt RC,et al, "Linkage analysis of susceptibility to ozone-induced lung inflammation in inbred mice.",*Nat Genet*,vol.17,n.4,1997,pp.475-8.
34. Otto-Knapp R, Jurgovsky K et al,"Antioxidative enzymes in human nasal mucosa after exposure to ozone. Possible role of GSTM1 deficiency,*Inflamm Res*,vol,52,n.2,2003,pp.51-5.
35. Reddy P, Naidoo RN et al," GSTM1, GSTP1, and NQO1 Polymorphisms and Susceptibility to Atopy and Airway Hyperresponsiveness among South African Schoolchildren", *Lung*, 2010.
36. Romieu I, Moreno-Macias H, "Gene by environment interaction and ambient air pollution", *Proc Am Thorac Soc*, vol.7, 2010, pp:116-22.
37. David GL, Romieu I,"Infant mortality and air pollution: modifying effect by social class." *J Occup Environ Med*,vol.46,n.12,2004,pp. 1210-6.
38. Cozier YC, Palmer JR, Rosenberg L," Comparison of methods for collection of DNA samples by mail in the Black Women's Health Study",*Ann Epidemiol*, vol.14,n.2,2004,pp.117-22.
39. Sigurdson AJ, Ha M, Cosentino M et al," Long-term storage and recovery of buccal cell DNA from treated cards", *Cancer Epidemiol Biomarkers Prev*,vol. 15,n.4,2006,pp.816-9.
40. Rogers NL, Cole SA et al." New saliva DNA collection method compared to buccal cell collection techniques for epidemiological studies", *Am J Hum Biol*,vol.19n.3,2007,pp.319-26.
41. Dawes C, "Estimates, from salivary analyses, of the turnover time of the oral mucosal epithelium in humans and the number of bacteria in an edentulous mouth, *Arch Oral Biol*, vol.48, 2003, pp:329-36..
42. Quinque D, Kittler R, Kayser M,et al," Evaluation of saliva as a source of human DNA for population and association studies.",*Anal Biochem*, vol.353,n.2,2006,pp.272-7.

43. Walker AH, Najarian D, et al," Collection of genomic DNA by buccal swabs for polymerase chain reaction-based biomarker assays", *Environ Health Perspect*, vol.107, 1999,pp:517-20.
44. Lum A, Le Marchand L, "A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies", *Cancer Epidemiol Biomarkers Prev*, vol. 7, 1998, pp:719-24.
45. Helene C et al" DNA elution from buccal cells stored on Whatman FTA Classic Cards using a modified methanol fixation method", *Biotechnique*, vol.46,2009,pp.309-311.
46. Dellinger AE, Saw SM, "Comparative analyses of seven algorithms for copy number variant identification from single nucleotide polymorphism arrays", *Nucleic Acids Res*, vol. 9, 2010.
47. Hawkins TL, Detter JC, Richardson PM," Whole genome amplification--applications and advances", *Curr Opin Biotechnol*,vol. 13,n.1,2002,pp.65-7.
48. Jasmine F, Ahsan H et al," Whole-genome amplification enables accurate genotyping for microarray-based high-density single nucleotide polymorphism array" *Cancer Epidemiol Biomarkers Prev*, vol.17,n.12,2008,pp.3499-508.
49. Thompson MD, Bowen RA et al," Whole genome amplification of buccal cell DNA: genotyping concordance before and after multiple displacement amplification",*Clin Chem Lab Med*,vol.43,n.2,2005,pp.157-62.
50. Holbrook JF, Stabley D, Sol-Church K Exploring whole genome amplification as a DNA recovery tool for molecular genetic studies. *J Biomol Tech*, vol.16,n.2,2005,pp.125-33.
51. Murthy KK, Mahboubi VS, Santiago A, et al. Assessment of multiple displacement amplification for polymorphism discovery and haplotype determination at a highly polymorphic locus, MC1R. *Hum Mutat*,vol.26,n.2,2005,pp.145-52.
52. Tranah GJ, Lescault PJ et al," Multiple displacement amplification prior to single nucleotide polymorphism genotyping in epidemiologic studies",*Biotechnol Lett*,vol.25,n.13,2003,pp.1031-6.
53. Dickson PA, Montgomery GW, Henders A, Campbell MJ, Martin NG, James MR.Queensland Institute of Medical Research, Brisbane, 4029 Australia. Evaluation of multiple displacement amplification in a 5 cM STR genome-wide scan.*Nucleic Acids Res*,vol.33,n.13,2005.
54. Sun G, Kaushal R et al," Whole-genome amplification: relative efficiencies of the current methods",*Leg Med (Tokyo)*,vol.7,n.5,2005,pp.279-86.
55. Jasmine F, Ahsan H, Andrulis IL et al," Whole-genome amplification enables accurate genotyping for microarray-based high-density single nucleotide polymorphism array",*Cancer Epidemiol Biomarkers Prev*, vol.12,2008, pp:3499-508.

56. Paez JG, Lin M, Beroukhir R et al, "Genome coverage and sequence fidelity of phi29 polymerase-based multiple strand displacement whole genome amplification", *Nucleic Acids Res*, vol.32, 2004, pp:e71.
57. Iafrate AJ, Feuk L, "Detection of large-scale variation in the human genome", *Nat Genet*, vol.9, 2004, pp:949-51 .
58. Randi A. Paynte," Accuracy of Multiplexed Illumina Platform-Based Single-Nucleotide Polymorphism Genotyping Compared between Genomic and Whole Genome Amplified DNA Collected from Multiple Sources," *Cancer Epidemiol Biomarkers Prev*, vol 12, 2006.
59. Corneveaux JJ, Kruer MC, "SNP-based chromosomal copy number ascertainment following multiple displacement whole-genome amplification", *Biotechniques*, vol.42, 2007, pp:77-83.
60. Bergen AW, Qi Y et al," Effects of DNA mass on multiple displacement whole genome amplification and genotyping performance", *BMC Biotechnol*, vol.5,n.24,2005.
61. Levieil K, Olarte M,et al, "Genotyping accuracy for whole-genome amplification of DNA from buccal epithelial cells", *Twin Res*, vol.7, 2004, pp:482-4. cells.
62. Leanza SM, Burk RD,et al, "Whole genome amplification of DNA extracted from hair samples: potential for use in molecular epidemiologic studies", *Cancer Detect Prev*, vol.31, 2007, pp:480-8.

9. PRODOTTI SCIENTIFICI DELL'ATTIVITA' DI DOTTORATO

Congenital Hyperinsulinism of infancy:a dominant KATP channel in an Italian family; Abstract LWPES/ESPE 8th Joint Meeting Global Care in Pediatric Endocrinology, *Hormone Research*:72(suppl 3), 2009.

- Iperinsulinismo congenito:mutazione dominante nel canale K+ATP-dipendente in una famiglia italiana;XVII Congresso nazionale- Società di endocrinologia e diabetologia pediatrica (SIEDP), 2009.

- L'esperienza italiana sull'iperinsulinismo congenito dell'infanzia: il registro nazionale e analisi mutazionale di *abcc8* e *kcnj11*: XIII congresso nazionale- Società Italiana di Genetica Umana (SIGU), Firenze 2010.